THIS OPINION WAS NOT WRITTEN FOR PUBLICATION

The opinion in support of the decision being entered today (1) was not written for publication in a law journal and (2) is not binding precedent of the Board.

Paper No. 11

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

KENNETH A. BARTON and MICHAEL J. MILLER,
(U.S. Application 07/827,906)

(Barton),

or

DAVID A. FISCHHOFF and FREDERICK J. PERLAK,
(U.S. Application 08/434,105)

Junior Party (Fischhoff),

v.

MICHAEL J. ADANG, THOMAS A. ROCHELEAU,
DONALD J. MERLO and ELIZABETH E. MURRAY,
(U.S. Patent 5,380,831)

Senior Party (Adang)

Interference 103,781

Interference 103,781

DECISION AND ORDER ON PRELIMINARY
AND MISCELLANEOUS MOTIONS AND REQUESTS

TEDDY S. GRON, Administrative Patent Judge.

1. Background

September 26, 1983 - Michael J. Adang and John D. Kemp filed U.S. Application 06/535,354, entitled "Insect Resistant Plants" (assignment to Agrigenetics Research Associates Ltd. recorded September 20, 1983; assignment to Lubrizol Genetics Inc. recorded May 29 1986), now abandoned.

April 4, 1986 - Michael J. Adang and John D. Kemp filed U.S. Application 06/848,733, entitled "Insect Resistant Plants" (assignment to Lubrizol Genetics Inc. recorded June 25, 1986), now abandoned, which is claimed to be a continuation-in-part of U.S. Application 06/535,354, filed September 26, 1983.

September 9, 1988 - Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray filed U.S. Application 07/242,482, entitled "Synthetic Insecticidal Crystal Protein Gene" (assignment to Lubrizol Genetics Inc. recorded October 24, 1988), now abandoned, which is claimed to be a continuation-in-part of U.S. Application 06/848,733, filed April 4, 1986, which is claimed to be a continuation-in-part of U.S. Application 06/535,354, filed September 26, 1983.

February 24, 1989 - David A. Fischhoff and Frederick J. Perlak filed U.S. Application 07/315,355, entitled "Synthetic Plant Genes And Method For Preparation" (assignment to Monsanto Company recorded February 24, 1989), now abandoned.

August 7, 1989 - Kenneth A. Barton and Michael J. Miller filed U.S. Application 07/390,561, entitled "Expression of Genes In Plants" (assignment to Agracetus recorded August 7, 1989; assignment to Monsanto Company recorded October 15, 1996), now abandoned.

February 12, 1990 - David A. Fischhoff and Frederick J. Perlak filed U.S. Application 07/476,661, entitled "Synthetic Plant Genes And Method For Preparation" (assignment to Monsanto Company recorded February 12, 1990), now abandoned, which is claimed to be a continuation-in-part of U.S. Application 07/315,355, filed February 24, 1989.

January 28, 1992 - Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray filed U.S. Application 07/827,844, entitled "Synthetic Insecticidal Crystal Protein Gene" (assignment to Mycogen Plant Science, Inc., recorded April 1, 1993), now abandoned, which is claimed to be a continuation-in-part of U.S. Application 07/242,482, filed September 9, 1988, which is claimed to be a continuation-in-part of U.S. Application 06/848,733, filed April 4, 1986, which is

claimed to be a continuation-in-part of U.S. Application 06/535,354, filed September 26, 1983.

January 30, 1992 - Kenneth A. Barton and Michael J. Miller filed involved U.S. Application 07/827,906, entitled "Improved Expression of Genes In Plants" (assignment to Monsanto Company recorded October 15, 1996; assignment to Monsanto Technology LLC recorded June 13, 2001), which is claimed to be a continuation of U.S. Application 07/390,561, filed August 7, 1989.

October 9, 1992 - David A. Fischhoff and Frederick J. Perlak filed U.S. Application 07/959,506, entitled "Synthetic Plant Genes" (assignment to Monsanto Technology LLC recorded June 13, 2001), which is claimed to be a continuation of U.S. Application 07/476,661, filed February 12, 1990, which is claimed to be a continuation-in-part of U.S. Application 07/315,355, filed February 24, 1989.

May 3, 1993 - Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray filed U.S. Application 08/057,191, entitled "Synthetic Insecticidal Crystal Protein Gene" (assignment to Mycogen Plant Science, Inc., recorded April 1, 1993), which is claimed to be a continuation of U.S. Application 07/827,844, filed January 28, 1992, which is claimed to be a continuation-in-part of U.S. Application 07/242,482, filed September 9, 1988, which is claimed to be a continuation-

in-part of U.S. Application 06/848,733, filed April 4, 1986, which is claimed to be a continuation-in-part of U.S. Application 06/535,354, filed September 26, 1983.

January 6, 1995 - Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray filed U.S. Application 08/369,839, entitled "Synthetic Insecticidal Crystal Protein Gene," which is claimed to be a division of 08/057,191, filed May 3, 1993, which is claimed to be a continuation of U.S. Application 07/827,844, filed January 28, 1992, which is claimed to be a continuation-in-part of U.S. Application 07/242,482, filed September 9, 1988, which is claimed to be a continuation-in-part of U.S. Application 06/848,733, filed April 4, 1986, which is claimed to be a continuation-in-part of U.S. Application 06/535,354, filed September 26, 1983.

January 6, 1995 - Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray filed U.S. Application 08/369,835, entitled "Synthetic Insecticidal Crystal Protein Gene" (assignment to Mycogen Plant Science, Inc., recorded April 1, 1993), which is claimed to be a continuation-in-part of 08/057,191, filed May 3, 1993, which is claimed to be a continuation of U.S. Application 07/827,844, filed January 28, 1992, which is claimed to be a continuation-in-part of U.S. Application 07/242,482, filed September 9, 1988, which is claimed

to be a continuation-in-part of U.S. Application 06/848,733, filed April 4, 1986, which is claimed to be a continuation-in-part of U.S. Application 06/535,354, filed September 26, 1983.

January 10, 1995 - involved U.S. Patent 5,380,831 issued from Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray, U.S. Application 08/057,191, filed May 3, 1993.

May 3, 1995 - David A. Fischhoff and Frederick J. Perlak, filed involved U.S. Application 08/434,105 entitled "Synthetic Plant Genes And Method For Preparation" (assignment to Monsanto Technology LLC recorded June 13, 2001), which is claimed to be a divisional of U.S. Application 07/959,506, filed October 9, 1992, which is claimed to be a continuation of U.S. Application 07/476,661, filed February 12, 1990, which is claimed to be a continuation-in-part of U.S. Application 07/315,355, filed February 24, 1989.

May 19, 1996 - U.S. Patent 5,500,365 issued from David A. Fischhoff and Frederick J. Perlak, U.S. Application 07/959,506, filed October 9, 1992.

August 29, 1996 - Michael J. Adang and Elizabeth E. Murray filed U.S. Application 08/705,438, entitled "Synthetic Insecticidal Crystal Protein Gene Having A Modified Frequency Of Codon Usage" (assigned to Mycogen Plant Science, Inc.), which is

claimed to be a division of U.S. Application 08/369,835, filed January 6, 1995 (now U.S. Patent 5,567,600), which is claimed to be a continuation-in-part of 08/057,191, filed May 3, 1993 (now U.S. Patent 5,380,831), which is claimed to be a continuation of U.S. Application 07/827,844, filed January 28, 1992, which is claimed to be a continuation-in-part of U.S. Application 07/242,482, filed September 9, 1988.

August 29, 1996 - Michael J. Adang and Elizabeth E. Murray filed U.S. Application 08/704,966, entitled "Transgenic Plants Comprising Synthetic Insecticidal Crystal Protein Gene Having A Modified Frequency Of Codon Usage" (assigned to Mycogen Plant Science, Inc.), which is claimed to be a division of U.S. Application 08/369,839, filed January 6, 1995 (now U.S. Patent 5,567,862), which is claimed to be a continuation-in-part of 08/057,191, filed May 3, 1993 (now U.S. Patent 5,380,831), which is claimed to be a continuation of U.S. Application 07/827,844, filed January 28, 1992, which is claimed to be a continuation-in-part of U.S. Application 07/242,482, filed September 9, 1988.

October 22, 1996 - U.S. Patent 5,567,600 issued from Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray, U.S. Application 08/369,835, filed June 6, 1995, including a disclaimer of the terminal portion of the

patent extending beyond the expiration date of U.S. Patent 5,380,831.

October 22, 1996 - U.S. Patent 5,567,862 issued from Michael J. Adang, Thomas A Rocheleau, Donald J. Merlo and Elizabeth E. Murray, U.S. Application 08/369,839, filed June 6, 1995.

November 7, 1996 - Interference 103,781 was declared essentially as follows (Paper No. 2):

JUNIOR PARTY APPLICATION

Named Inventors:	Kenneth A. Barton and Michael J. Miller
Application:	Application 07/827,906, filed January 30, 1992
Title:	Improved Expression of Genes in Plants
Assignee:	None (assignment to Monsanto Company recorded October 15, 1996; assignment to Monsanto Technology LLC recorded June 13, 2001)
Accorded benefit for the purpose of priority of:	Application 07/390,561, filed August 7, 1989

JUNIOR PARTY APPLICATION

Named Inventors:	David A. Fischhoff and Frederick J. Perlak
Application:	Application 08/434,105, filed May 3, 1995
Title:	Synthetic Plant Genes and Method for Preparation

Assignee:	None (assignment to Monsanto Technology LLC recorded June 13, 2001)
Accorded benefit for the purpose of priority of:	Application 07/959,506, filed October 9, 1992, now U.S. Patent 5,500,365, issued March 3, 1996; Application 07/476,661, filed February 12, 1990, now abandoned; and Application 07/315,355, filed February 24, 1989, now abandoned

SENIOR PARTY PATENT

Named Inventors:	Michael J. Adang, Thomas A. Rocheleau, Donald J. Merlo and Elizabeth E. Murray
Application:	Application 08/057,191, filed May 3, 1993, now U.S. Patent 5,380,831, issued January 10, 1995
Title:	Synthetic Insecticidal Crystal Protein Gene
Assignee:	Mycogen Plant Science, Inc. (Paper No. 13)
Accorded benefit for the purpose of priority of:	Applications 07/827,844, filed January 28, 1992, now abandoned, and Application 07/242,482, filed September 9, 1988, now abandoned

Count 1

A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

a) analyzing the coding sequence of a gene derived from Bacillus thuringiensis which encodes an insecticidal protein toxin, and modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence, or

b) analyzing the coding sequence of a gene derived from Bacillus thuringiensis which encodes an insecticidal protein toxin, and modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence and fewer plant polyadenylation signals than said coding sequence.

The claims of the parties which correspond to this count are:

Barton et al.: Claims 1-4, 7, and 15-22

Fischhoff et al.: Claims 3, 5, and 39-43

Adang et al.: Claims 1-14.

December 12, 1996 - An administrative patent judge (APJ) entered an Order to Show Cause stating (Paper No. 11, pp. 1-2, bridging para.):

In view of the common ownership by Monsanto Company of the Barton application and the Fischhoff application, the junior party Barton is ordered to show cause why judgement should not be entered against him within 30 days from the date of this order. Monsanto Company, as the assignee of both Barton and Fischhoff, may name the prior inventor in response to this order. Cf. M.P.E.P. 2302.

January 17, 1997 - The APJ ordered Monsanto Company "to name the prior inventor of count 1 . . . . In the event Monsanto makes no election, judgment will be entered against junior party Barton" (Paper No. 29, p. 3).

February 3, 1997 - Junior Party Barton et al. (hereafter Barton) petitioned the Commissioner under 37 CFR § 1.644(a)(1)

to reverse or postpone the APJ's January 17, 1997 order (Paper No. 35).

March 26, 1997 - Barton's February 3, 1997, petition was denied (Paper No. 38).

June 19, 1997 - The Board of Patent Appeals and Interferences (hereafter Board) entered the following judgment (Paper No. 53):

> Whereas Monsanto, the common assignee of the Barton et al. and Fischhoff et al. applications has named the party Fischhoff et al. as the prior inventor of count 1, pursuant to 37 CFR 1.602(a) and 1.610(e) judgement is hereby entered against Barton et al. as to the subject matter of count 1. Accordingly, Kenneth A. Barton and Michael J. Miller are not entitled to a patent containing Claims 1-4, 7, and 15-22 corresponding to count 1. The interference will continue as Fischhoff et al. v. Adang et al.

June 27, 1997 - Barton filed notice under 35 U.S.C. §§ 141 and 142 of appeal to the U.S. Court of Appeals for the Federal Circuit from the judgment of the Board of Patent Appeals and Interferences entered June 17, 1997 (Paper No. 55).

February 5, 1998 - The U.S. District Court for the District of Delaware entered a judgment (Mycogen Plant Science, Inc. v. Monsanto Co., No. 96-505 (D. Del. Feb. 5, 1998)) in a suit brought by Mycogen Plant Science, Inc., and Agrigenetics Inc. against Monsanto Co., DeKalb Genetics Corp., and Delta and Pine Land Co. for infringement of two Mycogen patents (Adang et al., U.S. Patent 5,567,862, entitled "Synthetic Insecticidal Crystal

Protein Gene," which issued October 22, 1996, from U.S. Application 08/369,839, filed January 6, 1995, and Adang et al., U.S. Patent 5,567,600, entitled "Synthetic Insecticidal Crystal Protein Gene," which issued October 22, 1996, from U.S. Application 08/369,835, filed January 6, 1995). A jury rendered a verdict that (1) defendants' products did not literally infringe any of the contested claims of Mycogen's patents, and (2) all of the contested claims of Mycogen's patents are invalid because Monsanto invented the subject matter thereof before the priority dates of Mycogen's patents. See the Procedural History in Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 1320-1321, 58 USPQ2d 1030, 1033-1034 (Fed. Cir. 2001)(Paper No. 125).

December 9, 1998 - The Court of Appeals for the Federal Circuit reversed the Board's June 19, 1997 judgment and remanded (Paper No. 124). Barton v. Adang, 162 F.3d 1140, 49 USPQ2d 1128 (Fed. Cir. 1998)(Paper No. 118, Exhibit A).

September 8, 1999 - The U.S. District Court for the District of Delaware entered a revised order (Paper No. 125, Exh. H) and opinion ruling on post-trial motions (Paper No. 125, Exh. I) (Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. 2d 199 (D. Del. 1999)) which "granted Mycogen's motion for JMOL that Monsanto's processes and resulting products infringed . . . the

'600 and '862 patents." Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1320, 58 USPQ2d at 1034 (Paper No. 146).

> The district court also denied Mycogen's motion for JMOL requesting the court to set aside the jury's finding of anticipation due to prior invention by Monsanto.
>
> The district court granted Monsanto's motion for JMOL holding that the claims of the '600 and '862 patents were invalid for lack of enablement pursuant to 35 U.S.C. § 112. . . . .

Id. at 1321, 58 USPQ2d at 1034 (Paper No. 146).

November 10, 1999 - In a suit brought by Mycogen Plant Science, Inc. and Agrigenetics Inc. against Monsanto Company for infringement of plaintiff's patent (Adang et al., U.S. Patent 5,380,831, issued January 10, 1995, from U.S. Application 08/057,191, filed May 3, 1993), the U.S. District Court for the Southern District of California entered an order (Mycogen Plant Sci., Inc. v. Monsanto Co., No. 95-CV-653 (S.D. Cal. Nov. 10, 1999)(Paper No. 127, Exh. A) which granted defendant's motion for summary judgment that Claims 1-12 of Mycogen's '831 patent are invalid under 35 U.S.C. § 102(g) and/or § 103 because Monsanto invented the subject matter thereof before Mycogen, as determined by the U.S. District Court for the District of Delaware in Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp.2d 199 (D. Del. 1999), which was affirmed in Mycogen Plant Sci., Inc. v. Monsanto Inc., 243 F.3d 1316, 58 USPQ2d 1030 (Fed. Cir. 2001), and denied defendant's motion for summary judgment that the

contested claims of Mycogen's '831 patent are invalid for noncompliance with the enablement requirement of the first paragraph of 35 U.S.C. § 112 as moot (Paper No. 127, Exh. A).

January 18, 2000 - U.S. patent 6,015,891 issued from Michael J. Adang and Elizabeth E. Murray, U.S. Application 08/705,438, filed August 29, 1996, subject to disclaimers of its term extending beyond the statutory expiration dates of Adang et al., U.S. Patent 5,567,600, issued October 22, 1996, and Adang et al., U.S. Patent 5,380,831, issued January 10, 1995.[1]

---

[1] But for prior issued patents identified in the terminal disclaimers entered as Paper No. 23 in U.S. patent 6,015,891 and Paper No. 25 in U.S. Patent 6,013,523, both disclaimers read:

> The owner of 100% interest in the instant application, Mycogen Plant Science, Inc., hereby disclaims, except as provided below, the terminal part of the statutory term of any patent granted on the instant application, which would extend beyond the expiration date of the full statutory term defined in 35 U.S.C. 154 to 156 and 173 for U.S. Patent No . . . issued on . . . .
>
> In making the above disclaimer, the owner does not disclaim the terminal part of any patent granted on the instant application that would extend to the expiration date of the full statutory term as defined in 35 U.S.C. 154 to 156 and 173 of U.S. Patent No . . . in the event that any of the . . . patents: expires for failure to pay a maintenance fee, is held unenforceable, is found invalid by a court of competent jurisdiction, is statutorily disclaimed in whole or terminally disclaimed under 37 CFR 1.321, has all claims cancelled by a reexamination certificate, is reissued, or is in any manner terminated prior to the expiration of its full statutory term.

January 18, 2000 - U.S. patent 6,013,523 issued from Michael J. Adang and Elizabeth E. Murray, U.S. Application 08/704,966, filed August 29, 1996, subject to a disclaimer of its term extending beyond the statutory expiration date of Adang et al., U.S. Patent 5,567,862, issued October 22, 1996.

March 12, 2001 - On appeal from the decision of the U.S. District Court for the District of Delaware in Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp.2d 199 (D. Del. 1999), the U.S. Court of Appeals for the Federal Circuit:

> . . . affirm[ed] the verdict of noninfringement based on patent invalidity due to prior invention pursuant to 35 U.S.C. § 102(g). This makes it unnecessary to address the finding of lack of enablement pursuant to 35 U.S.C. § 112.

Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1320, 58 USPQ2d at 1033 (Paper No. 146).

May 30, 2001 - On appeal from the decision of the U.S. District Court for the Southern District of California in Mycogen Plant Sci., Inc. v. Monsanto Co., No. 95-CV-653 (S.D. Cal. Nov. 10, 1999)(Paper No. 127, Exh. A), the U.S. Court of Appeals for the Federal Circuit affirmed-in-part, reversed-in-part, and remanded. Mycogen Plant Sci., Inc. v. Monsanto Co., 252 F.3d 1306, 1306, 58 USPQ2d 1891, 1892-1893 (Fed. Cir. 2001). The Federal Circuit concluded at 1309, 58 USPQ2d at 1893, that:

> . . . the district court improperly resolved disputed questions of material fact pertaining to the issue of

> prior invention, and we therefore reverse the court's ruling on summary judgment that the '831 patent is invalid under 35 U.S.C. § 102(g). We decline to affirm the summary judgment of invalidity on the alternative ground of non-enablement, as urged by Monsanto, but leave to the district court the task of determining in the first instance whether there is a genuine issue of material fact as to enablement based on its assessment of the evidence presented to it in the summary judgment proceeding.

Id. at 1310, 58 USPQ2d at 1894, the Federal Circuit explained:

> We agree with the district court that collateral estoppel requires the court to conclude that Monsanto reduced the invention [claimed in the Mycogen's '831 patent] to practice before Mycogen, and that collateral estoppel does not resolve the question whether Mycogen was the first to conceive and then was diligent during the critical period. On the merits of the summary judgment question, however, we do not agree that Monsanto has met its burden of showing that there are no issues of material fact regarding whether Mycogen was the first to conceive the invention and then diligently reduce it to practice.

2. Outstanding requests and motions

A. Adang's Preliminary Motion 1 (Paper No. 45)

By Adang's Preliminary Motion 1 (Paper No. 45), Adang moves under 37 CFR § 1.633(c)(1) to redefine the interfering subject matter by substituting Proposed Substitute Count 2 for existing Count 1.

Fischhoff's Opposition (Paper No. 58)
Adang's Reply (Paper No. 96)

B. Adang's Preliminary Motion 2 (Paper No. 46)

By Adang's Preliminary Motion 2 (Paper No. 46), Adang moves under 37 CFR § 1.633(f) for benefit of the January 28, 1992,

filing date of U.S. Application 07/827,844, and the September 9, 1988 filing date of U.S. Application 07/242,482, for Proposed Substitute Count 2 of Adang's Preliminary Motion 1 (Paper No. 45).

Fischhoff's response (Paper No. 59)
Adang's Reply (Paper No. 97)

C. Adang's Contingent Preliminary Motion 3 (Paper No. 47)

By Adang's Contingent Preliminary Motion 3 (Paper No. 47), Adang moves under 37 CFR § 1.633(a) for judgment that Claims 3, 5, and 39-43 of Fischhoff's involved U.S. Application 08/434,105, filed May 3, 1995, all Fischhoff's claims designated as corresponding to the interference count, are unpatentable under 35 U.S.C. § 102(g) over at least one claim of Barton's U.S. Application 07/827,906, filed January 30, 1992, designated as corresponding to the count, or under 35 U.S.C. § 103 in view of prior art including at least one claim of Barton's U.S. Application 07/827,906, filed January 30, 1992, designated as corresponding to the count. The motion presumes that party Barton has been determined to be, or Monsanto Technology LLC, the assignee of both of Fischhoff's and Barton's involved applications, has elected to designate Barton as, first to invent the subject matter defined by current Count 1 as between parties Fischhoff and Barton (Paper No. 47, pp. 1-2). In conjunction therewith, Adang contingently moves for permission to seek

deposition and documentary discovery relevant to the determination and/or election of Barton as first to invent the subject matter defined by the count as between Fischhoff and Barton (Paper No. 47, p. 2).

Fischhoff's Opposition (Paper No. 62)
Adang's Reply (Paper No. 98)

D. Fischhoff's First 37 CFR 1.642 Request (Paper No. 78)

Fischhoff requests that the APJ exercise his discretion under 37 CFR § 1.642 to add commonly assigned Adang et al., U.S. Patent 5,567,600 (Fischhoff Exhibit 37 (FX 37)), issued October 22, 1996, to this interference, designate all twenty-four claims thereof as corresponding to the count, and set an additional preliminary motion period for the parties to file motions relative to the newly added patent (Paper No. 78, p. 1).

Adang's Opposition (Paper No. 63)
Fischhoff's Reply (Paper No. 102)

E. Fischhoff's Second 37 CFR 1.642 Request (Paper No. 79)

Fischhoff requests that the APJ exercise his discretion under 37 CFR § 1.642 to add commonly assigned Adang et al., U.S. Patent 5,567,862 (FX 3), issued October 22, 1996, to this interference, designate all twenty-four claims thereof as corresponding to the count, and set an additional preliminary motion period for the parties to file motions relative to the newly added patent (Paper No. 79, p. 1).

Adang's Opposition (Paper No. 64)
Fischhoff's Reply (Paper No. 103)

F. Fischhoff's Preliminary Motion 3 (Paper No. 80)

By Fischhoff's Preliminary Motion 3 (Paper No. 80), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 (FX 11), issued January 10, 1995, are unpatentable under 35 U.S.C. § 112, second paragraph (Paper No. 80, p. 1).

Adang's Opposition (Paper No. 65)
Fischhoff's Reply (Paper No. 104)

G. Fischhoff's Preliminary Motion 4 (Paper No. 81)

By Fischhoff's Preliminary Motion 4 (Paper No. 81), Fischhoff moves under 37 CFR § 1.633(c)(1) to redefine the subject matter of this interference by substituting any one of Fischhoff's Proposed Counts 2, 3 and 4 for original Count 1 (Paper No. 81, p. 1).

Adang's Opposition (Paper No. 69)
Fischhoff's Reply (Paper No. 108)

H. Fischhoff's Preliminary Motion 5 (Paper No. 82)

By Fischhoff's Preliminary Motion 5 (Paper No. 82), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 (FX 11), issued January 10, 1995, are unpatentable under 35 U.S.C. § 112, first paragraph (enablement requirement)(Paper N. 82, p. 1).

Adang's Opposition (Paper No. 66)
Fischhoff's Reply (Paper No. 105)

I. Adang's 37 CFR § 635 Motion For Order Implementing the CAFC Decision In Barton v. Adang (Paper No. 116)

Adang moves under 37 CFR § 1.635 (Paper No. 116) for an order implementing the decision in Barton v. Adang, 162 F.3d 1140, 49 USPQ2d 1128 (Fed. Cir. 1998)(Paper No. 118, Exhibit A).

Joint Opposition Fischhoff and Barton (Paper No. 119)
Adang's Reply (Paper No. 121)

J. Fischhoff's Preliminary Motion 6 (Paper No. 83)

By Fischhoff's Preliminary Motion 6 (Paper No. 83), Fischhoff moves under 37 CFR § 1.633(c)(2) to redefine the subject matter of the interference by adding proposed Claims 44 and 45 to Fischhoff's involved U.S. Application 08/434,105, filed May 3, 1995 (Paper No. 84), and designating the new claims as corresponding to the count (Paper 83, pa. 1).

K. Fischhoff's Preliminary Motion 7 (Paper No. 85)

By Fischhoff's Preliminary Motion 7 (Paper No. 85), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 (FX 11), issued January 10, 1995, are unpatentable under 35 U.S.C. § 102 or § 103 (Paper No. 85, p. 1).

Adang's Opposition (Paper No. 68)
Fischhoff's Reply (Paper No. 107)

L. Fischhoff's Preliminary Motion 8 (Paper No. 86)

By Fischhoff's Preliminary Motion 8 (Paper No. 86), Fischhoff moves under 37 CFR § 1.633(f) to be accorded benefit of the October 9, 1992, filing date of Fischhoff's U.S. Application 07/959,506; the February 12, 1990, filing date of Fischhoff's U.S. Application 07/476,661; and the February 24, 1989, filing date of U.S. Application 07/315,355, for Fischhoff's Proposed Counts 2, 3, and 4 (Fischhoff's Preliminary Motion 4, Paper No. 81)(Paper No. 86, p. 1).

M. Fischhoff's Preliminary Motion 9 (Paper No. 87)

By Fischhoff's Preliminary Motion 9 (Paper No. 87), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang et al., U.S. Patent 5,380,831, issued January 10, 1995, are unpatentable under 35 U.S.C. § 112, first paragraph (written description requirement)(Paper No. 87, p. 1).

Adang's Opposition (Paper No. 67)
Fischhoff's Reply (Paper No. 106)

N. Fischhoff's Preliminary Motion 10 (Paper No. 88)

By Fischhoff's Preliminary Motion No. 10, Fischhoff moves under 37 CFR § 1.633(c)(4) to redefine the subject matter of the interference by designating (1) Claims 41-43 of Fischhoff's involved U.S. Application 08/434,105, filed May 3, 1995, and (2) Claims 13-14 of Adang's U.S. Patent 5,380,831 (FX 11),

issued January 10, 1995, as not corresponding to the count (Paper No. 88, p. 1).

Adang's Opposition (Paper No. 70)
Fischhoff's Reply (Paper No. 109)

O. Fischhoff's Preliminary Motion 11 (Paper No. 89)

By Fischhoff's Preliminary Motion 11 (Paper No. 89), Fischhoff moves under 37 CFR § 1.633(c)(2), contingent on granting Fischhoff's Second 37 CFR § 642 Request (Paper No. 79), to redefine the subject matter of the interference by adding proposed Claims 46 to Fischhoff's involved U.S. Application 08/434,105, filed May 3, 1995 (Paper No. 90), and designating the new claim as corresponding to the count (Paper No. 89, p. 1).

P. Fischhoff's Preliminary Motion 12 (Paper No. 60)

By Fischhoff's Preliminary Motion 12 (Paper No. 60), Fischhoff moves, contingent upon the granting of Adang's Preliminary Motion 1 (Paper No. 45), for benefit under 37 CFR § 1.633(f) of the October 9, 1992, filing date of Fischhoff's U.S. Application 07/959,506, the February 12, 1990, filing date of Fischhoff's U.S. Application 07/476,661, and the February 25, 1989, filing date of Fischhoff's U.S. Application 07/315,355, for subject matter defined by Adang's Proposed Substitute Count 2.

Q. Fischhoff's 37 CFR § 1.641(a) Request (Paper No. 110)

Fischhoff requests that the APJ exercise his discretion under 37 CFR § 1.641(a), notify the parties that Claims 1-12 of

Adang's U.S. Patent 5,380,831, issued January 10, 1995, appear to be unpatentable under 35 U.S.C. § 112, first paragraph (best mode requirement), and set a time period for the parties to take testimony and present related evidence and argument (Paper No. 110, p. 2, para. I).

Adang's Response (Paper No. 111)

R. Fischhoff's 37 CFR § 1.635 Motion For Temporary Stay of Interference Under 37 CFR § 1.645(d) (Paper No. 118)

Fischhoff moves under 37 CFR § 1.635 for an order temporarily staying the interference proceedings under 37 CFR § 1.645(d) in anticipation of an impending decision of the U.S. District Court for the Southern District of California "on a motion for summary judgment that the claims of . . . Adang's . . . U.S. Patent No. 5,380,831 are invalid under 35 U.S.C. § 102(g) because of prior invention by . . . Fischhoff . . ." (Paper No. 118, p. 2, para. 1).

Adang's Opposition (Paper No. 120)
Fischhoff's Reply (Paper No. 122)
Fischhoff's Supplement To Motion (Paper No. 125)
Adang's Opposition To Supplement (Paper No. 126)

S. Fischhoff's 37 CFR § 1.635 Motion For Temporary Stay of Interference Under 37 CFR § 1.645(d) (Paper No. 127)

Fischhoff moves under 37 CFR § 1.635 for an order temporarily staying the interference proceedings under 37 CFR § 1.645(d) pending a decision on appeal to the Federal Circuit of

a decision of the U.S. District Court for the Southern District of California granting Monsanto's motion for summary judgment that claims of Adang's U.S. Patent 5,380,831 are invalid under 35 U.S.C. § 102(g) over the prior invention of Fischhoff (Paper No. 127, p. 2, para. I).

Adang's Opposition (Paper No. 128)
Fischhoff's Reply (Paper No. 132)

3. Decisions on Outstanding Requests and Motions

"A party filing a motion has the burden of proof to show that it is entitled to the relief sought in the motion. 37 CFR § 1.637(a).

A. Adang's Preliminary Motion 1 (Paper No. 45)

Adang moves under 37 CFR § 1.633(c)(1) to redefine the interfering subject matter by replacing existing Count 1 with Proposed Substitute Count 2. Existing Count 1 reads (Paper No. 45, p. 3; Paper No. 2):

Count 1

> A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:
>
> a) analyzing the coding sequence of a gene derived from Bacillus thuringiensis which encodes an insecticidal protein toxin, and modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence, or
>
> b) analyzing the coding sequence of a gene derived from Bacillus thuringiensis which encodes an insecticidal

protein toxin, and modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence and fewer plant polyadenylation signals than said coding sequence.

Proposed Substitute Count 2 reads (Paper No. 45, p. 2):

Proposed Substitute Count 2

A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes an insecticidal protein toxin, and

modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence;

or

A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes an insecticidal protein toxin,

modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence, and

further comprising the step of modifying a portion of said coding sequence to eliminate plant polyadenylation signals.

Adang submits Proposed Substitute Count 2 to eliminate the possibility that "the analyzing step in part b) of the current

count could be incorrectly interpreted as analyzing only for the presence of polyadenylation signals" (Paper No. 45, p. 8, last complete sentence). Adang explains (Paper No. 45, p. 8, second full para.):

> Unfortunately, part b) of the current count, which uses the language of Fischhoff's claim 39, recites both the claim 1 step of "modifying" to yield a modified sequence which contains a greater number of codons preferred by the intended plant host and the claim 4 step of "modifying" to eliminate plant polyadenylation sequences as a single step. As a result, the single step of part b) of the current count could be incorrectly interpreted as analyzing for any characteristic that would give the combined result of having a greater number of codons preferred by the intended plant host and eliminating plant polyadenylation signals.

The steps of the current count are not open to the interpretation Adang contemplates for two reasons. First, the alternative process defined by part b) of the current count more particularly comprises the step of "modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence and fewer plant polyadenylation signals than said coding sequence" (emphasis added). When all the language of the current count is considered, the analyzing step of the process comprising part b) must comprise separate analyses of the coding sequence of a Bt gene both (1) for plant codon preferences and (2) for polyadenylation signals, and the modifying step of the process comprising part b) must comprise

separate modifications of a portion of the coding sequence to yield a modified sequence which contains (1) a greater number of codons preferred by the intended plant host than did said coding sequence, and (2) fewer polyadenylation signals than said coding sequence. Relative to the coding sequence of the Bt gene analyzed, reduction in the number of polyadenylation signals present in the coding sequence cannot simultaneously result in a greater number of codons preferred by the intended plant host relative to the number of codons preferred by the intended plant host present in the originally analyzed Bt coding sequence.

The above interpretation of the current count is consistent with the interpretations of substantially identical language in the claims of Adang's U.S. Patents 5,567,600 and 5,567,862, both issued October 22, 1995, and the parent thereof, Adang's involved U.S. Patent 5,380,831, issued January 10, 1995, by the U.S. District Court for the District of Delaware. See the district court's claim interpretation in Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp.2d 199, 215 (D. Del. 1999), affirmed on review by the Federal Circuit in Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 1326-1327, 58 USPQ2d 1030, 1038-1039 (Fed. Cir. 2001):

> Mycogen contests a portion of the district court's claim construction. Specifically, Mycogen contends that the district court's definition of the "greater number of codons preferred" language in independent claims 1,

2, 13 and 14 of the '600 patent is erroneous. However, the claim construction issue here relates to both the '600 and the '862 patent, as well as the original '831 parent patent, as all three patents contain claims that use the language disputed herein. Claim 1 of the '600 patent is representative, and it reads as follows:

> 1. A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:
>
> (a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;
>
> (b) modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in the codon positions II and III in a region between plant polyadenylation signals in said coding sequence;
>
> (c) inserting said modified sequence into the genome of a plant cell; and
>
> (d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said synthetic Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin.

'600 patent, col. 31, lines 37-57 (emphasis added).

In Mycogen, the district court held that:

> The phrase "greater number of codons preferred," is satisfied where the newly-created synthetic gene has a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host,

> and where the synthetic gene has an overall distribution of codon usage that is closer to that of the intended plant host.

61 F.Supp.2d at 215. . . . .

Ultimately, the Federal Circuit said, Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1329-1330, 58 USPQ2d at 1041:

> [T]runcation could not satisfy the district court's definition of the claim 1 "greater number" definition, because it would not result in a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host. . . . .
>
> . . . Thus, the district court's claim construction regarding the "greater number of codons preferred" limitation was correct.

The parties present different views of the clarity of the language of current Count 1 relative to the clarity of the language of Proposed Substitute Count 2 and the scope of the subject matter encompassed by each count. However, the issues raised are more appropriately considered in the context of the parties' motions for judgment on the patentability of their claims designated as corresponding to the interference count under 35 U.S.C. § 112, second paragraph. Suffice it to say, the reasons Adang presents for redefining the interfering subject matter by replacing current Count 1 with Proposed Substitute Count 2 are based on an interpretation of the language of current Count 1 which the Federal Circuit has determined is incorrect as a matter of law.

Adang has not shown that it is entitled to the relief requested. Accordingly, Adang's Preliminary Motion 1 (Paper No. 45) is DENIED.[2]

B. Adang's Preliminary Motion 2 (Paper No. 46)

Adang's Preliminary Motion 2 (Paper No. 46) under 37 CFR § 1.633(f) to be accorded benefit of the filing dates of U.S. Applications 07/827,844 and 07/242,482 for subject matter of Proposed Substitute Count 2 is contingent upon Adang's Preliminary Motion 1 (Paper No. 45) under 37 CFR § 1.633(c)(1) to redefine the subject matter of this interference by replacing existing Count 1 with Proposed Substitute Count 2 being granted. Since Adang's Preliminary Motion 1 has been denied, Adang's Preliminary Motion 2 is DISMISSED.

C. Adang's Contingent Preliminary Motion 3 (Paper No. 47)

Barton v. Adang, 162 F.3d 1140, 1146, 49 USPQ2d 1128, 1134 (Fed. Cir. 1998)(Paper No. 118, Exhibit A) held that Monsanto Company had shown good cause why (1) it should not have been

---

[2] Given that the Delaware district court interpreted the phrase "greater number of codons preferred" appearing in many, but not all, of the claims of Mycogen's U.S. Patents 5,567,600 and 5,567,862 (the phrase also appears in Claim 1 of Mycogen's U.S. Patent 5,380,831 and Count 1 of this interference recites the language of Claim 1 of U.S. Patent 5,380,831), the issue yet to be considered is whether the subject matter of the interference count should be redefined utilizing the language of Claim 11 of U.S. Patent 5,380,831. This issue is raised by Fischhoff's Preliminary Motion 4 (Paper No. 81).

forced to elect the prior inventor of Count 1 as between Fischhoff and Barton before the subject matter defined by the interference count is finalized, and (2) Barton should not have been prematurely dismissed as a party to this interference:

> [A]t the time that Monsanto was forced to make an election between the Barton et al. and the Fischhoff et al. applications, it was not clear what the content of the final count would be or what proofs on dates of conception and reduction to practice Adang et al. would seek to establish. Also, if the final count as decided by the Board excluded subject matter disclosed in Barton et al., but not Fischhoff et al., Monsanto loses arguably patentable subject matter by early dismissal of Barton et al.
>
> At this stage of the proceedings, Monsanto could not determine which application, either Barton et al. or Fischhoff et al., would be the best evidence to establish priority. We hold that Monsanto has shown "good cause" to continue the interference on both its applications until the preliminary motions to finalize the count are decided by the Board and discovery is complete.

Because Monsanto Company was improperly ordered to elect the prior inventor of Count 1 as between Barton and Fischhoff before the interference count was finalized, neither Fischhoff's nor Barton's claims designated as corresponding to Count 1 may be rejected under 35 U.S.C. § 102(g)/103 based on a forced election of the prior inventor of Count 1 by Monsanto Company at a time during the proceedings when "Monsanto could not determine which application, either Barton et al. or Fischhoff et al., would be the best evidence to establish priority." Id. Moreover, even if

Monsanto's forced election of the prior inventor of Count 1 as between Barton and Fischhoff could be basis for rejecting the nonelected party's claims under 35 U.S.C. § 102(g)/103, Monsanto's election of Fischhoff as the prior inventor of Count 1 most certainly does not support Adang's Contingent Preliminary Motion 3 under 37 CFR § 1.633(a) for a judgment that Fischhoff's claims designated as corresponding to Count 1 are unpatentable under 35 U.S.C. § 102(g)/103 over Barton's claims designated as corresponding to Count 1.

Because the Federal Circuit directed this interference proceeding to continue with both Fischhoff and Barton as parties "until the preliminary motions to finalize the count are decided by the Board and discovery is complete," Barton v. Adang, 162 F.3d at 1146, 49 USPQ2d at 1134, the discovery Adang requests is here, and is generally, inappropriate before the preliminary motion stage of the interference proceeding has been completed. Accordingly, Adang's Contingent Preliminary Motion 3 and the order for discovery Adang requests (Paper No. 47) are DENIED.

D. Fischhoff's First and Second 37 CFR § 1.642 Requests (Paper Nos. 78 & 79)

Fischhoff (Monsanto) asks the APJ to exercise his discretion under 37 CFR § 1.642 to (1) add Adang's (Mycogen's) U.S. Patent 5,567,600 (FX 37) and U.S. Patent 5,567,862 (FX 3), both issued October 22, 1996, to this interference, (2) designate all claims

of each patent as corresponding to the count, and (3) set an additional preliminary motion period for the parties to file motions relative to the newly added patents (Paper No. 78, p. 1; Paper No. 79, p. 1). However, on September 8, 1999, the U.S. District Court for the District of Delaware held that the claims of both Mycogen's U.S. Patent 5,567,600 (FX 37) and Mycogen's U.S. Patent 5,567,862 (FX 3) are invalid as unpatentable (1) under 35 U.S.C. § 102(g)/103 due to prior invention of the subject matter thereof by Monsanto, and (2) under 35 U.S.C. § 112, first paragraph, as based on a nonenabling disclosure. Mycogen Plant Science, Inc. v. Monsanto Co., 61 F. Supp.2d 199 (D. Del. 1999). The district court's decision that the claims of Mycogen's U.S. Patents 5,567,600 (FX 37) and 5,567,862 (FX 3) are invalid as unpatentable under 35 U.S.C. § 102(g)/103 was affirmed on appeal to the Federal Circuit. Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 1320, 58 USPQ2d 1030, 1034 (Fed. Cir. 2001)(Paper No. 146).

The parties to this interference have not discussed the Delaware district court's conclusions that the claims of Mycogen's U.S. Patent 5,567,600 (FX 37) and the claims of Mycogen's U.S. Patent 5,567,862 (FX 3) are invalid as unpatentable under 35 U.S.C. § 102(g)/103 because of prior invention by Monsanto or as unpatentable under 35 U.S.C. § 112,

first paragraph (nonenablement). Mycogen has neither shown nor argued that the Delaware district court's conclusions are incorrect. Moreover, the Federal Circuit affirmed the Delaware district court's holding that the claims of Mycogen's patents are invalid under 35 U.S.C. § 102(g) over Monsanto's prior invention. Accordingly, Fischhoff's requests to redefine the interfering subject matter by adding Mycogen's U.S. Patents 5,567,600 (FX 37) and 5,567,862 (FX 3) to this interference and designating the claims of those patents as corresponding to the count are DISMISSED.

There is a related matter. On January 18, 2000, U.S. Patent 6,015,891 (Appendix A) issued from Michael J. Adang and Elizabeth E. Murray, U.S. Application 08/705,438, filed August 29, 1996, including a terminal disclaimer of its term which extends beyond the statutory expiration dates of Mycogen's U.S. Patent 5,567,600 and U.S. Patent 5,380,831, issued January 10, 1995, and U.S. Patent 6,013,523 (Appendix B) issued from Michael J. Adang and Elizabeth E. Murray, U.S. Application 08/704,966, filed August 29, 1996, including a terminal disclaimer of its term which extends beyond the statutory expiration date of Mycogen's U.S. Patent 5,567,862. To completely resolve all patentability issues concerning what appears to be claims drawn to the "same patentable invention" as the claims presently designated as

corresponding to the count and priority of invention for the subject matter defined by the count as between the parties to this interference, a comparison of the representative method and gene claims presented in Mycogen's U.S. Patents 5,380,831, 5,567,600, 5,567,862, 6,013,523, and 6,015,891, is in order.

Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831 read:

> 1. A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:
>
> analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes an insecticidal protein toxin, and
>
> modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence.
>
> 11. A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:
>
> analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes an insecticidal protein toxin, and
>
> modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed.

Claim 1 of Mycogen's U.S. Patent 6,013,523 reads:

> 1. A descendant plant cell comprising a pesticidal protein toxin encoded by a synthetic Bacillus thuringiensis (B.t) gene, said cell produced by the process of:
>
> selecting a B.t. pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native B.t. gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant in which it is to be expressed than did the native B.t. coding sequence having at least about 10% of the nucleotides changed as compared to the native B.t. coding sequence;

obtaining a synthetic B.t. gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic B.t. gene into a plant cell;

culturing said cell to obtain descendant plant cells or plants comprising descendant plant cells, said descendant plant cells comprising said synthetic B.t. gene; and

establishing that said synthetic B.t. gene is expressed in said descendant plant cells.

Claims 1 and 4 of Mycogen's U.S. Patent 6,015,891:

1. A synthetic Bacillus thuringiensis (B.t.) gene which is expressed in descendant plant cells and encodes a pesticidal protein toxin, wherein said synthetic B.t. gene is produced by the process of:

selecting a B.t. pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native B.t. gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage

of the plant cell in which it is to be expressed than did the native B.t. coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native B.t. coding sequence;

obtaining a synthetic B.t. gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic B.t. gene into a plant cell;

culturing said cell to obtain descendant plant cells or plants comprising descendant plant cells, said descendant plant cells comprising said synthetic B.t. gene; and

establishing that said synthetic B.t. gene is expressed in said descendant plant cells.

4. A method of designing a synthetic Bacillus thuringiensis (B.t) gene which is expressed in descendant plant cells, comprising the steps of:

selecting a B.t. pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native B.t. gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native B.t. coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native B.t. coding sequence;

obtaining a synthetic B.t. gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic B.t. gene into a plant cell;

culturing said cell to obtain descendant plant cells, said descendant plant cells comprising said synthetic B.t. gene; and

establishing that said synthetic B.t. gene is expressed in said descendant plant cells.

Claims 1, 2, 7, 8, 13, 14, 19 and 20 of Mycogen's U.S. Patent 5,567,600 (FX 37), which were determined by the Delaware district court to be invalid, read:

1. A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence;

(c) inserting said modified sequence into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said synthetic Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin.

2. A DNA coding sequence produced by:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal

protein toxin; and

(b) modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence.

7. A method of designing a synthesis Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence;

(c) inserting said modified sequence into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said synthetic Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin.

8. A DNA coding sequence produced by:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin; and

(b) modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon

usage of the plant in which it is to be expressed than did said coding sequence, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence.

13. A method of designing a synthesis Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence prior to modification, and wherein said modification results in fewer occurrences of the sequence AATGAA in said modified sequence than in said coding sequence;

(c) inserting said modified sequence into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said synthesis Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin.

14. A DNA coding sequence produced by:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin; and

(b) modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence, and wherein said modification results in fewer occurrences of the sequence AATGAA in said modified sequence than in said coding sequence.

19. A method of designing a synthesis Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed than did said coding sequence prior to modification, and wherein said modification results in fewer occurrences of the sequence AATGAA in said modified sequence than in said coding sequence;

(c) inserting said modified sequence into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said synthetic Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin.

20. A DNA coding sequence produced by:

(a) analog the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin; and

(b) modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed than did said coding sequence, and wherein said modification results in fewer occurrences of the sequence AATGAA in said modified sequence than in said coding sequence.

Claims 1, 6, 7, 12, 13, 18, 19, and 24 of U.S. Patent 5,567,862 (FX 3), read:

1. A plant cell comprising a heterologous modified structural gene derived from a Bacillus thuringiensis gene encoding a pesticidal protein toxin, said plant cell produced by the steps of:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) a portion of said coding sequence to yield a modified structural gene which contains a greater number of codons preferred by said plant cell than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence;

(c) inserting said modified sequence into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

6. A method of producing a pesticidal protein comprising the steps of:

(a) introducing into a host plant cell a heterologous modified structural gene derived from a Bacillus thuringiensis gene wherein the DNA coding sequence of the Bacillus thuringiensis gene has been modified to contain a greater number of codons preferred by said plant cell than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence, and

(b) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene in the genome of said additional plant

cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin

7. A plant cell comprising a heterologous modified structural gene derived from a Bacillus thuringiensis gene encoding a pesticidal protein toxin, said plant cell produced by the steps of:

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of genes native to said plant cell than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence;

(c) inserting said modified structural gene into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified gene in the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

12. A method of producing a pesticidal protein comprising the steps of

(a) introducing into a host plant cell a heterologous modified structural gene derived from a Bacillus thuringiensis gene wherein the DNA coding sequence of the Bacillus thuringiensis gene has been modified to contain a frequency of codon usage that more closely resembles the frequency of codon usage of genes native to said plant cell than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in codon positions II and III in a region between plant polyadenylation signals in said coding sequence; and

(b) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene in the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

13. A plant cell comprising a heterologous modified structural gene derived from a Bacillus thuringiensis gene encoding a pesticidal protein toxin, said plant cell produced by the steps of

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said codon sequence to yield a modified structural gene which contains a greater number of codons preferred by the intended plant host than did said coding sequence prior to modification, and wherein said modification results in fewer occurrences of the sequence AATGAA in said modified structural gene than in said coding sequence;

(c) inserting said modified structural gene into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene in the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

18. A method of producing a pesticidal protein comprising the steps of

(a) introducing into a host plant cell a heterologous modified structural gene derived from a Bacillus thuringiensis gene wherein the DNA coding sequence of the Bacillus thuringiensis gene has been modified to contain a greater number of codons preferred by said plant cell than did said coding sequence prior to modification, and wherein the modification results in fewer occurrences of the sequence AATGAA in said modified structural gene than in said coding sequence; and

(b) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene in the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

19. A plant cell comprising a heterologous modified structural gene derived from a Bacillus thuringiensis gene encoding a pesticidal protein toxin, said plant cell produced by the steps of

(a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin;

(b) modifying a portion of said coding sequence to yield a modified structural gene which has a frequency of codon usage which more closely resembles the frequency of codon usage of genes native to said plant cell than did said coding sequence prior to modification, and wherein said modification results in fewer occurrences of the sequence AATGAA in said modified sequence than in said coding sequence;

(c) inserting said modified structural gene into the genome of a plant cell; and

(d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene in the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

24. A method of producing a pesticidal protein comprising the steps of

(a) introducing into a host plant cell a heterologous modified structural gene derived from a Bacillus thuringiensis gene wherein the DNA coding sequence of the Bacillus thuringiensis gene has been modified to contain a frequency of codon usage that more closely resembles the frequency of codon usage of genes native to said plant cell than did the coding

sequence prior to modification, and wherein the modification results in fewer occurrences of the sequence AATGAA in said modified structural gene than in said coding sequence; and

(b) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified structural gene in the genome of said additional plant cells, wherein said modified structural gene is expressed to produce a pesticidal protein toxin.

Claims 1, 4, and 11 of Mycogen's U.S. Patent 5,380,831 are directed to methods of designing a synthetic B.t. gene to be more highly expressed in a plant than a native B.t. gene, comprising:

(a) analyzing the coding sequence of a native B.t. gene encoding an insecticide, and

(b) modifying a portion of said analyzed coding sequence

(i) to contain a greater number of codons preferred by said plant than did the analyzed coding sequence, and/or

(ii) to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of said plant.

By reference to substantially the same language appearing in claims of Mycogen's U.S. Patent 5,467,600, the Federal Circuit indicated that the subject matter defined by the language of step (b)(i) above, which also appears in the process defined by Claim 1 of Mycogen's U.S. Patent 5,380,831, is either synonymous

with, of subsumed by, the language of step (b)(ii) above, which appears in the process defined by Claim 11 of Mycogen's U.S. Patent 5,380,831. In Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 1327, 58 USPQ2d 1030, 1038 (Fed. Cir. 2001), the Federal Circuit quoted the following holding of the Delaware district court in reference to Claims 1, 2, 13, and 14 of Mycogen's U.S. Patent 5,567,600:

> [T]he phrase "greater number of codons preferred," is satisfied where the newly-created synthetic gene has a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host, and where the synthetic gene has an overall distribution of codon usage that is closer to that of the intended plant host.

Mycogen Plant Science, Inc. v. Monsanto Co., 61 F. Supp.2d 199, 215 (D. Del. 1999).

On appeal, Mycogen contested the Delaware district court's claim interpretation. See Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1326, 58 USPQ2d at 1038:

> Mycogen contests a portion of the district court's claim construction. Specifically, Mycogen contends that the district court's definition of the "greater number of codons preferred" language in independent claims 1, 2, 13 and 14 of the '600 patent is erroneous. However, the claim construction issue here relates to both the '600 and the '862 patent, as well as the original '831 parent patent, as all three patents contain claims that use the language disputed herein.

The Federal Circuit stated, 243 F.3d at 1327, 58 USPQ2d at 1039:

> [T]he district court's claim construction defines a "preferred codon" to be any codon that brings the modified Bt gene's codon frequency closer to that of the intended plant host.

Mycogen responded, Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1328-1329, 58 USPQ2d at 1040:

> [T]he district court's definition of "greater number of codons preferred" cannot stand because it renders this phrase synonymous with the "frequency of codon usage" limitation in independent claims 7, 8, 19 and 20 of the '600 patent, thus rendering these two sets of claims identical.

The Federal Circuit held that "the district court's claim construction regarding the 'greater number of codons preferred' limitation was correct." Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1330, 58 USPQ2d at 1041.

The two-step methods claimed in Mycogen's U.S. 5,380,831 are broader in scope than the four-step methods claimed in each of Mycogen's U.S. Patent 5,567,600, U.S. Patent 5,567,862, U.S. Patent 6,013,523, and U.S. Patent 6,015,891, because the claimed two-step methods "of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants" are open to the additional steps of "inserting said modified sequence into the genome of a plant cell; and . . . maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said

synthetic Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin"; e.g., Claims 1 and 7 of Mycogen's U.S. Patent 5,567,600. Thus, the Delaware district court's conclusion that the four-step methods claimed in Mycogen's U.S. Patents 5,567,600 and 5,567,862, which include the two-step methods claimed in Mycogen's U.S. Patent 5,380,831, are unpatentable under 35 U.S.C. § 102(g)/103 and/or 35 U.S.C. § 112, first paragraph (enablement requirement), does not require a conclusion that the two-step methods claimed in Mycogen's U.S. Patent 5,380,831 are unpatentable for all the same reasons. The enablement problems may stem from the added method steps. Showings required to establish priority of invention for four-step processes generally exceed showings required to establish priority of invention for at least two steps of the four-step processes. Furthermore, the Delaware district court's conclusion that claims of Mycogen's U.S. Patent 5,567,600, which are drawn to a DNA coding sequence produced by the method of Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831, are unpatentable under 35 U.S.C. § 102(g)/103 and/or 35 U.S.C. § 112, first paragraph, does not require a conclusion that method Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831 also are unpatentable under 35 U.S.C. § 102(g)/103 and/or 35 U.S.C. § 112, first paragraph, for the same reasons.

However, the conclusion that the DNA coding sequences of Claims 2 and 8 of Mycogen's U.S. Patent 5,567,600, which are produced by the methods of Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831, are directed to the "same . . . invention" as the methods of Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831, and accordingly, the "same . . . invention" as Count 1 of this interference, is inescapable because the processes defined by Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831 and the processes which define the products-by-process of Claims 2 and 8 of Mycogen's U.S. Patent 5,567,600 are identical. Moreover, the four-step processes of designing a synthetic B.t. gene to be more highly expressed in plants of Claims 1 and 7 of Mycogen's U.S. Patent 5,567,600 would have been obvious under 35 U.S.C. § 103 to persons having ordinary skill in the art in view of the two-step processes of Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831 for designing a synthetic B.t. gene to be more highly expressed in plants, because the two additional steps of the four-step processes of the '600 patent no more than generally recite the utility described for the two-step processes of the '831 patent. See 37 CFR § 1.642 and 37 CFR § 1.601(n). The product-by-process claims of Mycogen's U.S. Patent 5,567,862 are directed to the "same . . . invention" as the process claims of Mycogen's U.S. Patent 5,567,600 again because the processes

defined by the product-by-process claims and the process claims of the respective patents are identical.

Nevertheless, even though the Delaware district court held that the claims of Mycogen's U.S. Patents 5,567,600 and 5,567,862 are invalid, the claims of Mycogen's U.S. Patents 6,013,523 and 6,015,891 are presumed valid. Therefore, questions of greater import to the present interference are whether the claims of the later issued patents, which are presumed valid, are directed to the same patentable invention as Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831 and Count 1, and accordingly, whether the subject matter of the count should be redefined by (1) adding Mycogen's U.S. Patent 6,015,891 to this interference and designating Claims 1-6 thereof as corresponding to the count, and (2) adding Mycogen's U.S. Patent 6,013,523 to this interference and designating Claims 1-4 thereof as corresponding to the count.

First, the terminal portion of Mycogen's U.S. Patent 6,015,891 which extends beyond the expiration dates of Mycogen's U.S. Patents 5,567,600 and 5,380,831 has been disclaimed by Mycogen, and the terminal portion of Mycogen's U.S. Patent 6,013,523 which extends beyond the expiration date of Mycogen's U.S. Patents 5,567,862 has been disclaimed by Mycogen. Terminal disclaimers are necessitated by an examiner's rejection of claims of a later-issued, commonly assigned patent for obviousness-type

double patenting of the subject matter claimed in an earlier-issued, unexpired patent.

Second, Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1325, 58 USPQ2d at 1037, instructs:

> The district court analyzed the infringement decision and stated "[f]or every one of [Monsanto's] accused genes and gene products, Mycogen presented evidence, detailed below, showing that a particular codon usage table was used to design each accused gene in a way that infringes the '600 and '862 patents. Defendants did not challenge this evidence. Mycogen, 61 F. Supp. 2d at 245.

Mycogen has established that the common analysis and modification steps of the processes claimed in Mycogen's U.S. Patent 5,567,600 and U.S. Patent 5,567,862 involve the steps of (I) "obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native B.t. gene," and (II) "using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native B.t. coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native B.t. coding sequence" (Mycogen's U.S. 6,015,891, Claims 1 and 4; Mycogen's U.S. 6,013,523, Claim 1). Method Claims 1 and 11 of Mycogen's U.S. 5,380,831 comprise analysis and modification steps

substantially identical to those recited in the method claims of Mycogen's U.S. Patents 5,567,600 and 5,567,862. Moreover, the specification of Mycogen's U.S. 5,380,831 teaches (col. 7, l. 1-28)(emphasis added):

> Frequency of preferred codon usage refers to the preference exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. To determine the frequency of usage of a particular codon in a gene, the number of occurrences of that codon in the gene is divided by the total number of occurrences of all codons specifying the same amino acid in the gene. Table 1, for example, gives the frequency of codon usage for Bt genes, which was obtained by analysis of four Bt genes whose sequences are publicly available. Similarly, the frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell. It is preferable that this analysis be limited to genes that are highly expressed by the host cell. Table 1 (see page 43), for example, gives the frequency of codon usage by highly expressed genes exhibited by dicotyledonous plants, and monocotyledonous plants. The dicot codon usage was calculated using 154 highly expressed coding sequences obtained from Genbank which are listed in Table 1. Monocot codon usage was calculated using 53 monocot nuclear gene coding sequences obtained from Genbank and listed in Table 1, located in Example 1.
>
> When synthesizing a gene for improved expression in a host cell it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

Mycogen's U.S. Patent 5,380,831 expressly states (Col. 10, l. 62, to Col. 11, l. 3)(emphasis added):

> For example, in preferred embodiments, the synthetic insecticidal protein is strongly expressed in dicot plants, e.g., tobacco, tomato, cotton, etc., and hence,

> a synthetic gene under these conditions is designed to incorporate to advantage codons used preferentially by highly expressed dicot proteins. In embodiments where enhanced expression of insecticidal protein is desired in a monocot, codons preferred by highly expressed monocot proteins (given in Table 1) are employed in designing the synthetic gene.

See also the disclosure of Mycogen's U.S. Patent 5,380,831 at column 11, lines 27-59, with regard to preferable avoidance of the CG dinucleotide in codon positions II and III, including references to prior art disclosing the "two codon choice indices for quantifying CG and AT doublet avoidance in codon positions II and III" (col. 11, l. 38-41) calculated for the plant data presented and tabulated in Table 2 (col. 11, l. 44-59). Claim 3 of Mycogen's U.S. Patent 5,380,831 explicitly recites "the step of modifying a portion of said coding sequence to yield CG and AT doublet avoidance indices which more closely resemble those of the intended plant host" (Col. 38, l. 38-41, Claim 3).

Third, Mycogen's U.S. Patent 5,380,831 includes the following statements (Col. 14, l. 12-66; citations omitted):

> The recombinant DNA molecule carrying a synthetic structural gene under promoter control can be introduced into plant tissue by any means known to those skilled in the art. The technique used for a given plant species or specific type of plant tissue depends on the known successful techniques. . . . Once introduced into the plant tissue, the expression of the structural gene may be assayed by any means known to the art, and expression may be measured as mRNA transcribed or as protein synthesized. Techniques are known for the in vitro culture of plant tissue, and in a number of cases, for regeneration of whole plants. Procedures for

> transferring the introduced expression complex to commercially useful cultivars are known to those skilled in the art.
>
> . . . Once plant cells expressing a synthetic insecticidal structural gene under control of a plant expressible promoter are obtained, plant tissues and whole plants are then reproduced by conventional means and the introduced genes can be transferred to other strains and cultivars by conventional plant breeding techniques.

Given that Mycogen has terminally disclaimed the unexpired term of Mycogen's U.S. Patent 6,015,891 which extends beyond the expiration dates of Mycogen's U.S. Patents 5,567,600 and 5,380,831 and the unexpired term of Mycogen's U.S. Patent 6,013,523 which extends beyond the expiration date of Mycogen's U.S. Patent 5,567,862 and, presuming that the claims of Mycogen's U.S. Patent 5,567,600, U.S. Patent 5,567,862, and U.S. Patent 5,380,831 are directed to the same invention as Count 1 of this interference, it seems reasonable to conclude that the subject matter claimed of Mycogen's U.S. Patents 6,015,891 and 6,013,523 would have been obvious to persons having ordinary skill in the art in view of prior art including Claims 1 and 11 of Mycogen's U.S. Patent 5,380,831. Accordingly, arguments can be made that the claims of Mycogen's U.S. Patents 6,015,891 and 6,013,523 are directed to the "same patentable invention" as the claims in Mycogen's U.S. Patent 5,380,831 and Count 1 of this interference. Therefore, the parties should consider whether Mycogen's U.S.

Patents 6,015,891 and 6,013,523 should be added to this interference.

E. Fischhoff's Preliminary Motion 4 (Paper No. 81)

By Fischhoff's Preliminary Motion 4 (Paper No. 81), Fischhoff moves under 37 CFR § 1.633(c)(1) to redefine the interference count by substituting Proposed Count 2, 3, or 4 for existing Count 1 (Paper No. 81, p. 1). The alternatively proposed counts read as follows (Paper No. 81, pp. 14-16):

Fischhoff's Proposed Count 2

Claim 1 of the '831 patent;[3]

- or -

Claim 3 of the '105 application.[4]

---

[3] 1. ['831 patent] A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:

analyzing the coding sequence of a gene derived from Bacillus thuringiensis which encodes an insecticidal protein toxin, and

modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence.

[4] 3. ['105 application] A method for modifying a wild-type structural gene sequence which encodes an insecticidal protein of Bacillus thuringiensis to enhance the expression of said protein in plants which comprises:

a) removing polyadenylation signals contains [sic] in said wild-type gene while retaining a sequence which encodes said protein; and

Fischhoff's Proposed Count 3

All claims of the party Adang et al. that survive the 37 CFR 1.633(a) and 37 CFR (c)(4) motions filed by the party Fischhoff et al.;

- or -

All claims of the party Fischhoff et al. (except claim 43) that survive any 37 CFR 1.633(a) and 37 CFR (c)(4) motions filed by the party Adang et al.

Fischhoff's Proposed Count 4

A. A method of preparing a gene derived from a Bacillus thuringiensis, said method comprising the steps of:

(a) selecting a group of genes that encode expressed proteins in a host plant;

(b) calculating the frequency of usage for codons within the structural coding region in said group of genes;

(c) selecting a gene derived from Bacillus thuringiensis, which encodes an insecticidal protein having an amino acid sequence;

(d) preparing a gene encoding the same amino acid sequence as the gene selected in step © but which has a structural coding region with:

(i) a frequency of usage for codons that closely approaches the frequency of usage for codons determined in step (b);

(e) inserting the gene obtained in step (d) into said host plant's genome; and

---

b) removing ATTTA sequences contained in said wild-type gene while retaining a sequence which encodes said protein.

(f) obtaining a transgenic plant with enhanced expression of a protein derived from a Bacillus thuringiensis;

- or -

B. A method of preparing a gene derived from a Bacillus thuringiensis, said method comprising the steps of:

(a) selecting a gene derived from a Bacillus thuringiensis, which encodes an insecticidal protein having an amino acid sequence:

(b) preparing a gene encoding the same amino acid sequence as the gene selected in step (a) but which has a structural coding region with:
(i) fewer ATTTA sequences,

(ii) fewer plant polyadenylation signals, and

(iii) lower A+T content

(c) inserting the gene obtained in step (b) into a host plant's genome; and

(d) obtaining a transgenic plant with enhanced expression of a protein derived from a Bacillus thuringiensis.

The phrase "greater number of codons preferred" of Claim 1 of Mycogen's U.S. 5,380,831 has been interpreted as follows by the Delaware district court in Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp.2d at 215:

> [T]he phrase "greater number of codons preferred," is satisfied where the newly-created synthetic gene has a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host, and where the synthetic gene has an overall distribution of codon usage that is closer to that of the intended plant host.

Given that interpretation, Fischhoff argues that existing Count 2 ought to be replaced by Fischhoff's Proposed Count 2, 3, or 4 because "Count 1 is narrower than the corresponding definition of gene modifications in Claim 3 of the '105 application" (Paper No. 81, p. 4, l. 1-3). In effect, Fischhoff argues that Fischhoff's Proposed Count 2, 3, or 4 should be substituted for existing Count 1 because the invention of Claim 3 of Fischhoff's involved U.S. Application 08/434,105 is not encompassed by, and does not correspond to, the interfering subject matter defined by existing Count 1. Whether presented in a motion under 37 CFR § 1.633(c)(4) to redefine the interfering subject matter by designating Claim 3 of the '105 application as not corresponding to the count or in a motion under 37 CFR § 1.633(c)(1) to redefine the interfering subject matter by substituting a count, the issue is the same. Presuming that Claim 3 of Fischhoff's U.S. Application 08/434,105 stands properly designated as corresponding to Count 1, the question to be asked is whether Fischhoff has provided sufficient reasons why the relief it requests should be granted?

Claim 3 of Fischhoff's involved U.S. Application 08/434,105 reads (Paper No. 81, p. 4):

> 3. ('105 Appl.) A method for modifying a wild-type structural gene sequence which encodes an insecticidal protein of Bacillus thuringiensis to enhance the expression of said protein in plants which comprises:

a) removing polyadenylation signals contained in said wild-type gene while retaining a sequence which encodes said protein; and

b) removing ATTTA sequences contained in said wild-type gene while retaining a sequence which encodes said protein.

On its face, neither the step of part a) nor the step of part b) of Claim 3 of the Fischhoff's '105 application appears to create a synthetic gene having a "higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host," the interpretation given the language "greater number of codons preferred" of Count 1 of this interference and Claim 1 of Mycogen's U.S. Patent 5,380,831 by the Delaware district court. Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 58 USPQ2d 1030 (Fed. Cir. 2001), affirming Mycogen Plant Sci.,, Inc. v. Monsanto Co., 61 F. Supp.2d at 215. However, Count 1 of this interference alternatively comprises the step of "modify a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence and fewer plant polyadenylation signals than said coding sequence" (Paper No. 2, Count 1b,; emphasis added).

The criteria for determining whether a claim should be designated as not corresponding to the count is presented in

37 CFR § 1.637(c)(4):

> A preliminary motion seeking to designate an application or patent claim as not corresponding to the count shall:
>
> (i) Identify the claim and the count.
>
> (ii) Show that the claim does not defined [sic] the same patentable invention as any other claim whose designation in the notice declaring the interference as corresponding to the count the party does not dispute.

Fischhoff does not dispute that Claim 5 of Fischhoff's U.S. Application 08/434,105, as amended November 29, 1995, stands properly designated as corresponding to Count 1. Claim 5 of the '105 application, as amended November 29, 1995, reads:

> 5. The method of Claim 3 further comprising the use of plant preferred sequences in the removal of the polyadenylation signals and ATTTA sequences.

Claim 5 of the '105 application, as originally filed, differed from Claim 5, as amended November 29, 1995, only in its dependency on original Claim 4. Claim 4 of the '105 application, as originally filed, defines the scope of the subject matter encompassed by Claim 5 designated as corresponding to Count 1. Original Claim 4 reads:

> 4. A method of Claim 3 further comprising the removal of self-complementary sequences and replacement of such sequences with nonself-complementary DNA comprising preferred codons while retaining a structural gene sequence encoding said protein.

The designation of Claim 5 of the '105 application as corresponding to Count 1 is not disputed by Fischhoff. Thus, the only question to be decided is whether Claim 3 of the '105 application defines the same patentable invention as Claim 5 of the '105 application.

Claim 3 of the '105 application is directed to the same patentable invention as Claim 5 of the '105 application if the subject matter defined by Claim 3 would have been obvious under 35 U.S.C. § 103 to a person having ordinary skill in the art in view of prior art assumed to include the subject matter of Claim 5 of the '105 application. 37 CFR § 1.601(n). If it is properly dependent on Claim 3 of the '105 application, method Claim 5 of the '105 application must be construed to incorporate by reference all the process limitations of Claim 3 of the '105 application to which it refers and set forth a further limitation thereof. 35 U.S.C. § 112, fourth paragraph. Neither Fischhoff, Adang, nor the examiner responsible for initiating this interference rejected the patentability of Claim 5 of the '105 application under 35 U.S.C. § 112, fourth paragraph. Therefore, the conclusion is inescapable that the subject matter defined by independent Claim 3 of the '105 application would have been prima facie obvious to a person having ordinary skill in the art in view of the subject matter defined by dependent Claim 5 of the

'105 application which incorporates all the limitations of independent Claim 3 of the '105 application upon which it depends. Accordingly, Fischhoff's justification for redefining the subject matter of the interference by substituting substitute Count 2, 3, or 4 for Count 1 is undermined. Thus, Fischhoff has not met its burden to show it is entitled to the relief sought.

Mycogen also suggests that Monsanto's motion to redefine the interfering subject matter by substituting proposed Count 2, 3, or 4 for existing Count 1 is designed primarily to expand the scope of the subject matter defined by the interference count to include methods of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants comprising the steps of analyzing the coding sequence of a gene derived from B.t. which encodes an insecticidal protein toxin, and modifying a portion of said coding sequence to yield a modified sequence which contains either (1) a greater number of codons preferred by the intended plant than did said coding sequence, (2) fewer plant polyadenylation signals than did said coding sequence, or (3) fewer ATTTA sequences than did said coding sequence. Whether or not Monsanto's aim is to expand the scope of the interfering subject matter, and concomitantly, the scope of the evidence upon which Mycogen and Monsanto may rely to show priority with respect to the same patentable invention claimed by the parties to this

interference, Monsanto's motion may be warranted if the broader scope of the interfering subject matter defined by the substitute count better defines the scope of the patentable invention contested by the parties to this interference.

Does any one of Fischhoff's Proposed Counts 2, 3, and 4 better define the scope of the interfering subject matter for which priority of invention is contested by the parties to this interference than existing Count 1? If so, then Fischhoff's motion to redefine the subject matter claimed by substituting that proposed count for existing Count 1 should be granted.

Proposed Count 4 alternatively defines methods of preparing a gene derived from B.t. which includes extraneous steps upon which the parties do not rely for patentability and language and/or terms whose meanings remain subject to interpretation; e.g., "calculating the frequency of usage for codons within the structural coding region in said group of genes" (Proposed Count 4A(b)) and "lower A+T content" (Proposed Count 4B(b)(iii)). Proposed Count 3 creates more uncertainty than certainty as to the scope of the interfering subject matter. The full scope of the subject matter defined by Proposed Count 3 must be finally resolved before the parties can determine what evidence is relevant to their respective cases for priority of invention.

On the other hand, Proposed Count 2, which is directed alternatively to Claim 1 of Mycogen's U.S. Patent 5,380,831 and Claim 3 of Monsanto's U.S. Application 08/434,104, is not only well defined, but it broadly encompasses claimed subject matter which the parties consider best representative of their inventions. Although Claim 1 of Mycogen's U.S. Patent 5,380,831 and Claim 3 of Monsanto's U.S. Application 08/434,104 do not themselves appear to be directed to the same patentable invention, Claim 4 of Mycogen's U.S. Patent 5,380,831 and Claim 5 of Monsanto's U.S. Application 08/434,104, which are designated as corresponding to Count 1 (in-part a transcription of Claim 1 of Mycogen's U.S. Patent 5,380,831), clearly define the same patentable invention. See 37 CFR § 1.601(j):

> An interference-in-fact exists when at least one claim of a party that is designated to correspond to the count and at least one claim of an opponent which is designated to correspond to the count define the same patentable invention.

Nevertheless, given the Delaware district court's interpretation of language in claims of Mycogen's U.S. Patent 5,567,600 common to language in Claim 1 of Mycogen's U.S. Patent 5,380,831 in Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp.2d 199, 215 (D. Del. 1999), Fischhoff's Proposed Count 2 is not broad enough to encompass all claims that are patentable over the prior art and correspond to the count. See 37 CFR §§ 1.601(f) and 606.

Claims 11 and 12 of Mycogen's U.S. 5,380,831 are presumed to be patentable over the prior art and, on their face, appear to be broader in scope than the subject matter defined by either Claim 1 of Mycogen's U.S. Patent 5,380,831 or Claim 3 of Monsanto's U.S. Application 08/434,104. Accordingly, to the extent Fischhoff moves under 37 CFR § 1.633(c)(1) to redefine the interfering subject matter by substituting any one of proposed Counts 2, 3, and 4 for existing Count 1 (Paper No. 81), the motion is DENIED.

However, Fischhoff alternatively invites the APJ to substitute a count of its own construction which is not narrower in scope than any application claim that is patentable over the prior art which is designated as corresponding to the count or any patent claim designated to correspond to the count. The invitation is accepted.

It is ORDERED that Interference 103,781 is redeclared as:

JUNIOR PARTY APPLICATION (Fischhoff)

Named Inventors:	Kenneth A. Barton and Michael J. Miller
Application:	Application 07/827,906, filed January 30, 1992
Title:	Improved Expression of Genes in Plants
Assignee:	None (assignment to Monsanto Company recorded October 15, 1996; assignment to Monsanto Technology LLC recorded June 13, 2001)

Interference 103,781

Accorded benefit for the purpose of priority of:	Application 07/390,561, filed August 7, 1989
	- or -
Named Inventors:	David A. Fischhoff and Frederick J. Perlak
Application:	Application 08/434,105, filed May 3, 1995
Title:	Synthetic Plant Genes and Method for Preparation
Assignee:	None (assignment to Monsanto Technology LLC recorded June 13, 2001)
Accorded benefit for the purpose of priority of:	Application 07/959,506, filed October 9, 1992, now U.S. Patent 5,500,365, issued March 3, 1996; Application 07/476,661, filed February 12, 1990, now abandoned; and Application 07/315,355, filed February 24, 1989, now abandoned

SENIOR PARTY PATENT (Adang)

Named Inventors:	Michael J. Adang, Thomas A. Rocheleau, Donald J. Merlo and Elizabeth E. Murray
Application:	Application 08/057,191, filed May 3, 1993, now U.S. Patent 5,380,831, issued January 10, 1995
Title:	Synthetic Insecticidal Crystal Protein Gene
Assignee:	Mycogen Plant Science, Inc. (Paper No. 13)

Accorded benefit for the purpose of priority of: Applications 07/827,844, filed January 28, 1992, now abandoned, and Application 07/242,482, filed September 9, 1988, now abandoned

Count 2

Any one of Claims 1-4, 7, and 15-22 of Barton et al.'s Application 07/827,906, filed January 30, 1992;

- or -

Any one of Claims 3, 5, and 39-43 of Fischhoff et al.'s Application 07/827,906, filed January 30, 1992;

- or -

Any one of Claims 1-14 of Adang et al.'s U.S. Patent 5,380,831, which issued January 10, 1995, from U.S. Application 08/057,191, filed May 3, 1993.

The claims of the parties which correspond to this count are:

Barton et al.: Claims 1-4, 7, and 15-22

Fischhoff et al.: Claims 3, 5, and 39-43

Adang et al.: Claims 1-14.

F. Fischhoff's Preliminary Motion 8 (Paper No. 86)

By Fischhoff's uncontested Preliminary Motion 8 (Paper No. 86), Fischhoff moves under 37 CFR § 1.633(f) to be accorded benefit of the October 9, 1992, filing date of Fischhoff's U.S. Application 07/959,506; the February 12, 1990, filing date of Fischhoff's U.S. Application 07/476,661; and the February 24, 1989, filing date of U.S. Application 07/315,355, for Fischhoff's

Proposed Counts 2, 3, and 4 (Fischhoff's Preliminary Motion 4, Paper No. 81)(Paper No. 86, p. 1). In that Fischhoff's Preliminary Motion 4 (Paper No. 81) has been denied for Proposed Counts 2, 3, and 4, this motion also is DISMISSED.

However, the declaration was redeclared in Paragraph E above with new Count 2 defining the interfering subject matter. For Count 2 of the redeclared Interference 103,781, Fischhoff has been accorded benefit for purposes of priority of the October 9, 1992, filing date of U.S. Application 07/959,506, now U.S. Patent 5,500,365, issued March 3, 1996; the February 12, 1990, filing date of U.S. Application 07/476,661, now abandoned; and the February 24, 1989, filing date of U.S. Application 07/315,355, now abandoned.

G. Fischhoff's Preliminary Motion 11 (Paper No. 89)

By Fischhoff's Preliminary 11 (Paper No. 89), Fischhoff moves under 37 CFR § 1.633(c)(2) to redefine the interfering subject matter by adding proposed Claim 46 to Fischhoff's U.S. Application 08/434,105, filed May 3, 1995 (Paper No. 90) and designating the new claim as corresponding to the count. The Motion is contingent upon granting Fischhoff's Second 37 CFR § 642 Request (Paper No. 79). Since Fischhoff's Second 37 CFR § 642 Request (Paper No. 79) has been dismissed, this motion also is DISMISSED.

H. Fischhoff's Preliminary Motion 12 (Paper No. 60)

By Fischhoff's Preliminary Motion 12 (Paper No. 60), Fischhoff moves under 37 CFR § 1.633(f) for benefit of the October 9, 1992, filing date of Fischhoff's U.S. Application 07/959,506; the February 12, 1990, filing date of Fischhoff's U.S. Application 07/476,661; and the February 25, 1989, filing date of Fischhoff's U.S. Application 07/315,355; for the subject matter defined by Adang's Proposed Substitute Count 2. Fischhoff's motion is contingent upon granting Adang's Preliminary Motion 1 (Paper No. 45). Adang's Preliminary Motion 1 (Paper No. 45) has been denied. Therefore, this motion is DISMISSED.

I. Adang's 37 CFR § 635 Motion For Order Implementing the CAFC Decision In Barton v. Adang (Paper No. 116)

In Part 3, Paragraph E, of this decision, it was ordered that this interference was redeclared, in effect implementing Barton v. Adang, 162 F.3d 1140, 49 USPQ2d 1128 (Fed. Cir. 1998) (Paper No. 118, Exh. A). Adang's unopposed Rule 635 motion is GRANTED.

J. Fischhoff's Preliminary Motion 6 (Paper No. 83)

By Fischhoff's Preliminary Motion 6 (Paper No. 83), Fischhoff moves under 37 CFR § 1.633(c)(2) to redefine the interfering subject matter by amending Fischhoff's U.S.

Application 08/434,105 to add new Claims 44 and 45 (Paper No. 84) and designating new Claims 44 and 45 as corresponding to the count, now new Count 2 (see the Order redeclaring Interference 103,781 in Part 3, paragraph E, of this decision). The motion is unopposed by Adang. Nevertheless, Fischhoff has the burden of proof to show that it is entitled to the relief sought in its motion. 37 CFR § 1.637(a). To show that it is entitled to the relief sought in a motion under 37 CFR § 1.633(c)(2), the movant must (i) "[p]ropose an amended or added claim[;]" (ii) "[s]how that the claim proposed to be amended or added defines the same patentable invention as the count[;]" and (iii) "[s]how the patentability to the applicant of each claim proposed to be amended or added to the disclosure of the application . . . ." Part (b) of Fischhoff's proposed method Claim 44 and DNA coding sequence Claim 45 reads (Paper No. 83, pp. 2-3; Paper No. 84):

> (b) preparing a gene encoding the same amino acid sequence as the gene selected in step (a) but which has a structural coding region with:
>
> (i) fewer ATTTA sequences,
> (ii) fewer plant polyadenylation signals, and
> (iii) lower A+T content[.]

Fischhoff merely presumes that the method of proposed Claim 44 and the DNA coding sequence of proposed Claim 45, both of which recite the additional step (b)(iii) of preparing a B.t. gene having a structural coding region with lower A+T content, would

have been obvious to persons having ordinary skill in the art in view of prior art including the method of modifying a wild-type structural gene which encodes an insecticidal protein of B.t. defined by Claim 3 of Fischhoff's U.S. Application 08/434,105. Claim 3 of Fischhoff's U.S. Application 08/434,105 recites the following steps:

> a) removing polyadenylation signals contained in said wild-type gene while retaining a sequence which encodes said protein; and
>
> b) removing ATTTA sequences contained in said wild-type gene while retaining a sequence which encodes said protein . . .[;]

Fischhoff has not satisfied its burden to show that the additional step of preparing a B.t. gene having a structural coding region with lower A+T content would have been obvious to persons having ordinary skill in the art in view of prior art teachings including Claim 3 of Fischhoff's U.S. application 08/434,105. Nor has Fischhoff shown that proposed Claims 44 and 45 including step b) would have been obvious to persons having ordinary skill in the art in view of prior art including any other claim designated as corresponding to Count 2. Thus, Fischhoff has not established that the claims proposed to be amended or added define the same patentable invention as the count.

Similarly, Fischhoff has not established that proposed Claims 44 and 45 are patentable to Fischhoff. No claim which recites step b)(iii) of proposed Claims 44 and 45 has previously been presented in Fischhoff's U.S. Application 08/434,105 for interpretation by the examiner. When considered in light of the supporting specification, step b)(iii) of preparing a gene encoding the same amino acid sequence as the gene selected in step (a) but which has a structural coding region with "lower A+T content" appears susceptible to a wide variety of interpretations. Thus, to satisfy its burden to show that proposed Claims 44 and 45 are patentable to Fischhoff, Fischhoff minimally should have shown that the meaning of the step b)(iii) would have been reasonably clear to persons having ordinary skill in the art in light of its supporting specification. Fischhoff has done no more than point to the same or substantially the same language in the specification. Accordingly, Fischhoff has not shown that it is entitled to the relief sought. Therefore, Fischhoff's Preliminary Motion 6 (Paper No. 83) under 37 CFR § 1.633(c)(2) to redefine the interfering subject matter by amending Fischhoff's U.S. Application 08/434,105 to add new Claims 44 and 45 is DENIED.

K. Fischhoff's Preliminary Motion 10 (Paper No. 88)

By Fischhoff's Preliminary Motion 10 (Paper No. 88), Fischhoff moves under 37 CFR § 1.66(c)(4) to redefine the interfering subject matter by designating (1) Claims 41-43 of Fischhoff's U.S. Application 08/434,105, and (2) Claims 13-14 of Adang's U.S. Patent 5,380,831 (FX 11), as not corresponding to the count. Adang opposes the motion (Paper No. 70).

Each of Claims 41-43 of Fischhoff's U.S. Application 08/434,105 is directed to "[a] modified chimeric gene comprising a promoter which functions in plant cells operably linked to a structural coding sequence and a 3'-nontranslated region comprising a polyadenylation signal which functions in plants to cause the addition of polyadenylate nucleotides to the 3' end of the RNA" (Paper No. 88, p. 2). For the Claim 41 chimeric gene, the structural coding sequence encodes an insecticidal protein derived from B.t. tenebrionis and comprises a structurally identified 1791 nucleotide sequence (Paper No. 88, p. 2). For the Claim 42 chimeric gene, the structural coding sequence encodes an insecticidal protein derived from B.t. entomocidus and comprises a structurally identified 3567 nucleotide sequence (Paper No. 88, p. 2). For the Claim 43 chimeric gene, the structural coding sequence encodes a B.t. P2 insecticidal protein

and comprises a structurally identified 1905 nucleotide sequence (Paper No. 88, p. 2).

Each of Claims 13 and 14 of Adang's U.S. Patent 5,380,831 is directed to "[a] synthetic gene comprising [a particular segment of] the DNA sequence presented in Fig. 1" (Paper No. 88, p. 1). For Claim 13, the DNA segment spans nucleotides 1-1793 of the DNA sequence of Fig. 1. For Claim 14, the DNA segment spans nucleotides 1-1833 of the DNA sequence of Fig. 1.

Claim 40 of Fischhoff's U.S. Application 08/434,105, whose designation as corresponding to Count 2 Fischhoff does not contest, is more generally directed to (Paper No. 88, p. 1):

> 40. A synthetic gene which is derived from a Bacillus thuringiensis insecticidal protein toxin gene and which is more highly expressed in plants, wherein the coding sequence of said synthetic gene is modified to contain:
>
> a) a greater number of codons preferred by the intended plant host than said insecticidal protein toxin gene; and
>
> b) fewer polyadenylation signal sequences than said insecticidal protein toxin gene.

Adang's U.S. Patent 5,380,831 does not claim synthetic genes defined by a non-specifically identified structural coding sequence. However, Claims 1 and 4 of Adang's U.S. Patent 6,015,891 (Appendix A) read:

> 1. A synthetic Bacillus thuringiensis (B.t.) gene which is expressed in descendant plant cells and encodes

a pesticidal protein toxin, wherein said synthetic B.t. gene is produced by the process of:

selecting a B.t. pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native B.t. gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native B.t. coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native B.t. coding sequence;

obtaining a synthetic B.t. gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic B.t. gene into a plant cell;

culturing said cell to obtain descendant plant cells or plants comprising descendant plant cells, said descendant plant cells comprising said synthetic B.t. gene; and

establishing that said synthetic B.t. gene is expressed in said descendant plant cells.

4. A method of designing a synthetic Bacillus thuringiensis (B.t) gene which is expressed in descendant plant cells, comprising the steps of:

selecting a B.t. pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native B.t. gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native B.t. coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native B.t. coding sequence;

obtaining a synthetic B.t. gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic B.t. gene into a plant cell;

culturing said cell to obtain descendant plant cells, said descendant plant cells comprising said synthetic B.t. gene; and

establishing that said synthetic B.t. gene is expressed in said descendant plant cells.

Fischhoff argues that Claims 41-43 of Fischhoff's U.S. Application 08/434,105 and Claims 13 and 14 of Adang's U.S. Patent 5,380,831 are directed to a "separate patentable invention" from any of Fischhoff's and Adang's claims more generally directed to methods of designing synthetic genes derived from B.t. and synthetic genes derived from B.t. (Paper No. 88). Fischhoff proffers that the specific structural DNA sequences defined by Fischhoff's Claims 41-43 and Adang's 13 and 14 would not have been obvious to persons having ordinary skill in the art in view of prior art teachings including the subject matter defined by any one or all of Fischhoff's and

Adang's more general claims designated as corresponding to Count 2. To the contrary, Adang argues that the subject matter defined by one or more of Fischhoff's and Adang's general claims designated as corresponding to Count 2 would have led persons having ordinary skill in the art to make and use any one of the specific synthetic structural gene sequences designed and/or encompassed by Fischhoff's and Adang's more general claims designated as corresponding to the count with reasonable expectation that the synthetic B.t. gene produced would be more highly expressed in plants as compared to the corresponding unmodified native B.t. gene. Decision on Fischhoff's Preliminary Motion 10 is DEFERRED TO FINAL HEARING for the following reasons.

First, the evidence upon which Fischhoff and Adang may wish to rely in presenting their respective cases in chief for priority of the invention with regard to the subject matter more generally defined by Count 2 of this interference is likely to include evidence of conception and/or reduction to practice of the more limited subject matter to which Claims 41-43 of Fischhoff's U.S. Application 08/434,105 and Claims 13 and 14 of Adang's U.S. Patent 5,380,831 are directed. Any evidence Fischhoff and Adang present for or against a conclusion that Claims 41-43 of its application and Claims 13-14 of Adang's patent are directed to separate patentable inventions from

any of the more general claims of the parties designated as corresponding to Count 2 is likely to affect the weight to be accorded evidence relative to specific structural gene sequences which Fischhoff or Adang may later present for the purpose of establishing that it first conceived and/or reduced to practice an invention encompassed by the more general scope of subject matter of Count 2.

Second, if the specific structural sequences of Claims 41-43 of Fischhoff's U.S. Application 08/434,105 and Claims 13 and 14 of Adang's U.S. Patent 5,380,831 are not directed to the same patentable invention as the generic synthetic gene sequences defined by Fischhoff's and Adang's more general claims designated as corresponding to Count 2, then the adequacy of the written descriptive support for, and conception of an invention of, Fischhoff's and Adang's remaining claims designated as corresponding to Count 2, which more generally define structural gene sequences, may be questioned. See Amgen Inc. v. Chugai Pharm. Co., 927 F.2d 1200, 18 USPQ2d 1016 (Fed. Cir.), cert. denied, 502 U.S. 856 (1991); Regents of the Univ. of Cal. v. Eli Lilly & Co., 119 F.3d 1559, 43 USPQ2d 1398 (Fed. Cir. 1997); and Enzo Biochem, Inc. v. Gen-Probe Inc., 285 F.3d 1-13, 62 USPQ2d 1289 (Fed. Cir. 2002), vacated-in-part, reversed-in-part,

and remanded On Petition For Rehearing Appeal 01-1230, decided July 15, 2002 (slip opinion).

Therefore, Fischhoff's Preliminary Motion 10 (Paper No. 88) is DEFERRED TO FINAL HEARING.

L. Fischhoff's Preliminary Motion 3 (Paper No. 80)

By Fischhoff's Preliminary Motion 3 (Paper No. 80), Fischhoff moves under 37 CFR § 1.633(a) for a judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 (FX 11) are unpatentable under 35 U.S.C. § 112, second paragraph. According to Fischhoff, the claims of Adang's involved patent do not particularly point out and distinctly claim the subject matter Adang regards as his invention for a variety of reasons considered separately below. It is essential that this motion be considered on the merits as early in the proceedings as possible, because questions related to the patentability of claimed subject matter under 35 U.S.C. § 112, first paragraph, § 102, and § 103, for example, may not be adequately considered until the full scope of the claimed subject matter is clear.

The language for which Claims 1-10 of Adang's patent are criticized first is the phrase "a modified sequence which contains a greater number of codons preferred by the intended plant host" (Paper No. 80, pp. 6-13) which appears in independent Claim 1 (FX 11). According to Fischhoff, "there is no definition

in the '831 patent of 'codons preferred by the intended plant host" (Paper No. 80, p. 6, last full para.). However, in Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 58 USPQ2d 1030 (Fed. Cir. 2001)(Paper No. 125), the Federal Circuit stated at 1326-1327, 58 USPQ2d at 1038-1039 (emphasis added):

> Mycogen contests a portion of the district court's claim construction. Specifically, Mycogen contends that the district court's definition of the "greater number of codons preferred" language in independent claims 1, 2, 13 and 14 of the '600 patent is erroneous. However, the claim construction issue here relates to both the '600 and the '862 patent, as well as the original '831 parent patent, as all three patents contain claims that use the language disputed herein. Claim 1 of the '600 patent is representative, and it reads as follows:
>
> > 1. A method of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants, comprising the steps of:
> >
> > (a) analyzing the coding sequence of a gene derived from a Bacillus thuringiensis which encodes a pesticidal protein toxin,
> >
> > (b) modifying a portion of said coding sequence to yield a modified sequence which contains a greater number of codons preferred by the intended plant host than did said coding sequence prior to modification, said modification comprising reducing the number codons having CG in the codon positions II and III in a region between plant polyadenylation signals in said coding sequence;
> >
> > (c) inserting said modified sequence into the genome of a plant cell; and

> (d) maintaining said plant cell under conditions suitable to allow replication of said plant cell to produce additional plant cells having said modified sequence in the genome of said additional plant cells, wherein said synthetic Bacillus thuringiensis gene is expressed to produce a pesticidal protein toxin.

'600 patent, col. 31, lines 37-57 (emphasis added).

In Mycogen, district court held that:

> [T]he phrase "greater number of codons preferred," is satisfied where the newly-created synthetic gene has a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host, and where the synthetic gene has an overall distribution of codon usage that is closer to that of the intended plant host.

61 F.Supp.2d at 215. Thus, the district court's claim construction defines a "preferred codon" to be any codon that brings the modified Bt gene's codon frequency closer to that of the intended plant host.

The Federal Circuit said, Mycogen Plant Science Inc. v. Monsanto Co., 243 F.3d at 1327, 58 USPQ2d at 1041:

> . . . Thus, the district court's claim construction regarding the "greater number of codons preferred" limitation was correct.

As indicated above, the '831 patent claim language "greater number of codons preferred," here criticized under 35 U.S.C. § 112, second paragraph, for clarity, has been construed by both the Delaware district court and the Court of Appeals for the Federal Circuit to mean that "the newly-created synthetic gene has a higher number of those codons whose frequency in the native

Bt gene was lower than their frequency in the intended plant host, and where the synthetic gene has an overall distribution of codon usage that is closer to that of the intended plant host" based on a common patent specification. Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1326-1327, 58 USPQ2d at 1038-1039. Therefore, Fischhoff's present motion under 37 CFR § 1.633(a) for judgment that Claims 1-10 of Adang's U.S. Patent 5,380,831 (FX 11) are unpatentable under 35 U.S.C. § 112, second paragraph, because the phrase "a modified sequence which contains a greater number of codons preferred by the intended plant host" (FX 11, Claim 1) does not particularly point out or distinctly claim the subject matter Adang regards as its invention must be denied. The scope of those claims of Adang's U.S. Patent 5,380,831 (FX 11) which are limited by that criticized language has been determined and is clear.

Second, Fischhoff argues that the phrases "the intended plant host" of Claim 1 and "the plant in which it is to be expressed" of Claim 11 of Adang's U.S. Patent 5,380,831 themselves render Claims 1-12 of Adang's U.S. Patent 5,380,831 vague and indefinite (Paper No. 80, pp. 14-18). The argument is based on Fischhoff's perception that skilled artisans would interpret the word "plant" in the two phrases to refer to "the whole plant," an interpretation which is inconsistent with

statements in the specification of Adang's U.S. Patent 5,380,831 (Paper No. 80, p. 14). For example, immediately following the "Frequency Of Codon Usage" listing in Adang's Table 1 (Adang's U.S. Patent 5,380,831, Table 1, col. 18), there is stated (emphasis added):

> 154 coding sequences of dicot nuclear genes were used to compile the codon usage table. The pooled dicot coding sequences, obtained from Genbank (release 55) or, when no Genbank file is specified, directly from the published source . . . .

Thus, the step of "modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of 'the plant in which it is to be expressed'" of Claim 11 of Adang's U.S. Patent 5,380,831 would appear to read on the step of "modifying a portion of said coding sequence to yield a modified sequence which has a frequence of codon usage which is more closely resembles the frequency of codon usage of [the nuclear genes] 'of the plant in which it is to be expressed'." According to Fischhoff, the frequency of codon usage for the whole plant is likely to differ markedly from the frequency of codon usage of the plant's nuclear genes. More particularly, Fischhoff notes from Adang's disclosure that "chloroplasts . . . have their own genome with different codon frequencies than are founds in nuclear genes" (Paper No. 80, p. 14, para. 2). See Adang's U.S.

Patent 5,380,831, col. 26, l. 27-49. In addition, Fischhoff argues that the frequency of codon usage varies considerably not only from one type of plant to another and from one type of tissue to another for the same plant, but also within each cell itself (Paper No. 80, p. 14).

Neither the Delaware district court nor the Court of Appeals for the Federal Circuit appear to have interpreted the term "plant" in the claims of Adang's patents. Presuming that the plain meaning of the term is "the whole plant," the fact that Adang's patent specification preferably modifies a portion of the coding sequence "to yield a modified sequence which contains a greater number of codons preferred by the intended plant host" (Claim 1 of Adang's U.S. Patent 5,380,831) based on the Frequency of Codon Usage Distribution Fractions compiled from 154 coding sequences of nuclear genes does not itself render the subject matter claimed vague and indefinite. Adang's specification teaches (Adang's U.S. Patent 5,380,831, col. 7, l. 11-16) (emphasis added):

> [T]he frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell. <u>It is preferable that this analysis be limited to genes that are highly expressed by the host cell</u>.

Again, Adang's patent teaches, "[T]o optimize the efficiency of translation, codons preferred in highly expressed proteins of

the host cell are utilized" (Adang's U.S. 5,380,831, col. 9, l. 52-54). Therefore, it is not surprising that Adang's patent specification discloses, as its preferred embodiment, modification of a portion of the coding sequence "to yield a modified sequence which contains a greater number of codons preferred by the intended plant host" (Claim 1 of Adang's U.S. Patent 5,380,831) based on the Frequency of Codon Usage Distribution Fractions compiled from 154 coding sequences of nuclear genes. Claim 11 of Adang's patent specification reflects a preference for "modifying a portion of said coding sequence to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed" (Claim 11 of Adang's U.S. Patent 5,380,831). Adang's patent specification teaches (Adang's U.S. 5,380,381, col. 26, l. 18-20 and 46-47):

> In general, the plant codon usage pattern more closely resembles that of man and other higher eukaryotes than unicellular organisms. . . .
>
> . . . . .
>
> . . . In general, the chloroplast codon profile more closely resembles that of unicellular organisms . . . .

Adang teaches that the codon bias of its nuclear genes best reflects the codon bias of a host plant, the whole plant. Adang teaches that the codon bias of genes which are highly expressed

in the host plant best reflect the codon bias of the whole plant. While the methods of Adang's patented claims modify a coding sequence to yield a modified sequence which contains a greater number of codons preferred in the intended plant host, Adang's specification teaches that the codon preference of the whole plant is best represented by the codon preferences of nuclear genes which are highly expressed in the host plant.

Third, while Adang's patent specification does not explicitly define "analyzing the coding sequence of a gene derived from Bacillus thuringiensis which encodes an insecticidal protein toxin" (Claims 1 and 11 of Adang's U.S. 5,380,831), it is clear that the native B.t. coding sequence must be analyzed to an extent necessary for one skilled in the art to modify a portion of the native B.t. coding sequence to yield a modified sequence which either contains a greater number of codons preferred by the intended plant host than did the native B.t. coding sequence analyzed or has a frequence of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed. There is nothing in the second paragraph of 35 U.S.C. § 112 which indicates that a broad step of a patentable method claim cannot be defined by its function as long as the claim language as a whole apprizes persons skilled in the art of the full scope of the method claimed. While Fischhoff prima

facie criticizes the term "analyzing" in the claimed methods as broad enough to encompass any kind of analysis which allows persons skilled in the art to modify a portion of the B.t. coding sequence as indicated, Fischhoff has not explained why the term itself does not satisfy 35 U.S.C. § 112, second paragraph, if the purpose of the analysis is clear (Paper No. 80, pp. 18-19). There is nothing inherently wrong with broad claim language.

Fourth, Fischhoff does not argue that the U.S. District Court for the District of Delaware, U.S. District Court for the Southern District of California, or the Court of Appeals for the Federal Circuit misinterpreted the methods of Claims 1-12 of Adang's U.S. Patent 5,308,831 comprising the steps of "analyzing . . ." and "modifying . . ." as further limited by the functional phrase "of designing a synthetic Bacillus thuringiensis gene to be more highly expressed in plants" (Paper No. 80, pp. 19-21). On its face, each of the methods of Claims 1 and 11 of Adang's U.S. Patent 5,380,831 comprises two steps which, upon performance thereof, are said to effectively design a synthetic Bacillus thuringiensis gene which is more highly expressed in a plant than an unmodified native B.t. gene is expressed in the same plant. Presuming the steps of the claimed methods adequately define the scope and content of the subject matter claimed, how persons skilled in the art would measure comparatively higher expression

of synthetic B.t. genes made by the methods of Claims 1-12 of Adang's patent appears to be immaterial to the patentability of the claims. Fischhoff has not shown that performance of the two steps of each of the claimed methods would not inherently design a synthetic B.t. gene which is more highly expressed in plants than the corresponding native B.t. gene.

Fifth, supported by declaratory evidence accompanied by data, Fischhoff argues that the phrase "codons preferred by the intended plant host" in Claim 1 of Adang's U.S. Patent 5,380,831 is vague and indefinite because the determinations whether codons are preferred by the intended plant host depend on the kind and number of genes used to determine the codons preferred by the plant host and the statistical differences associated with each determination (Paper No. 80, pp. 21-22). Even assuming that Fischhoff's argument has merit, Fischhoff has not shown that the scope of the claimed subject matter is vague, indefinite, and indeterminable. The Delaware district court held:

> [T]he phrase "greater number of codons preferred," is satisfied where the newly-created synthetic gene has a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host, and where the synthetic gene has an overall distribution of codon usage that is closer to that of the intended plant host.

Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp.2d 199, 215 (D. Del. 1999). On review, the Court of Appeals for the Federal

Circuit stated, "[T]he district court's claim construction regarding the 'greater number of codons preferred' limitation was correct." Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d at 1330, 58 USPQ2d at 1041.

Sixth, Fischhoff argues that the phrase "a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed" of Claims 11-12 of Adang's U.S. Patent 5,380,831 is itself indefinite (Paper No. 80, pp. 23-27). Initially, Fischhoff asks what the phrase "more closely resembles" means? However, the Federal Circuit indicated that the Delaware district court correctly construed the phase in context as follows:

> [I]t would . . . result in a higher number of those codons whose frequency in the native Bt gene was lower than their frequency in the intended plant host.

Id., 243 F.3d at 1329, 58 USPQ2d at 1041. Fischhoff has not explained why the Delaware district court's literal interpretation of the quoted claim language is unreasonable. Fischhoff acknowledges (Paper No. 80, p. 27, first full para.):

> Another reasonable interpretation of claims 11 and 12 is that any one amino acid's codon distribution, if changed in the right direction, causes the sequence to have a frequency of codon usage which more closely resembles the frequency of codon usage in the host. On that basis, Dr. Barry concluded that a prior art Bt fusion gene disclosed in DeGreve et al. (which contained a greater proportion of CCC's (proline) than the native Bt gene did) fully meets the literal language of the party Adang et al.'s claims . . . .

Fischhoff has not shown that the language of Claims 1-12 of Adang's U.S. Patent No. 5,380,831 is vague, indefinite, or indeterminable. Fischhoff has not shown that persons skilled in the art reasonably would not have understood the scope of the subject matter claimed in Adang's patent. Accordingly, Fischhoff's Preliminary Motion 3 (Paper No. 80) is DENIED.

M. Fischhoff's Preliminary Motion 5 (Paper No. 82)

By Fischhoff's Preliminary Motion 5 (Paper No. 82), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 are unpatentable under 35 U.S.C. § 112, first paragraph, as based on a specification which would not have enabled one skilled in the art to make and use the full scope of the invention claimed. This motion is DEFERRED TO FINAL HEARING.

Substantial portions of Fischhoff's arguments in support of its position that the subject matter defined by Claims 1-12 of Adang's U.S. Patent 5,380,831 is broader than the enabling disclosure are devoted to the confusion persons skilled in the art would have had trying to determine the scope and content of the claimed subject matter. As Fischhoff did in its Preliminary Motion 3 (Paper No. 80), Fischhoff argues that the phrases "to be more highly expressed in plants" in Claims 1-12 of Adang's patent (Paper No. 82, pp. 5-6), "which contains a greater number

of codons preferred by the intended plant host" in Claims 1-10 of Adang's patent (Paper No. 82, pp. 6-17), and "to yield a modified sequence which has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed" in Claims 11-12 of Adang's patent (Paper No. 82, pp. 17-25), are inadequately defined in the patent's specification. Because the phrases purport to be vague and indefinite, Fischhoff argues that persons skilled in the art would have been required to perform undue experimentation to determine how to modify a coding sequence derived from a native B.t. gene which encodes an insecticidal protein toxin to be more highly expressed in plants. Presumably then, since Fischhoff's Preliminary Motion 3 (Paper No. 80) has been denied, this motion also should be denied. Not necessarily so.

In its revised post-trial opinion of September 8, 1999 (Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp. 2d 199 (D. Del. 1999)), the Delaware district court considered the issue whether the claims of Adang's U.S. Patents 5,567,600 and 5,567,862 are invalid for lack of enablement (Paper No. 125, Attachment H, pp. 127-135). Much of the language of Claims 1-12 and the disclosure of the specification of Adang's U.S. Patent 5,380,831 upon which Fischhoff here bases its nonenablement arguments are common to the claims and specifications of Adang's

U.S. Patents 5,567,600 and 5,567,862. The Delaware district court held that the claims of Adang's U.S. Patents 5,567,600 and 5,567,862, which include further limitations, are invalid for lack of enablement (Paper No. 125, Attachment H, pp. 127-135). The court stated (Paper No. 125, Attachment H, pp. 134-135):

> Adang and Murray have identified a specific methodology for designing a gene by modifying codon sequences to reduce certain codons and have set out in their specification an example showing how they have implemented that methodology to design a synthetic Bt gene. In claiming their invention, they have identified the methodology, but have not identified which codons or how many should be removed. In light of the number of these codons in a gene, there are millions and even billions of ways to implement their methodology as claimed, almost all of which will not achieve the desired result. The example Adang and Murray have set out [in (sic)] the specification may show an example of how they have implemented their methodology, but neither the example nor the specifications provides guidance to those skilled in the art on how the methodology should be implemented. That is, the claims and specifications do not identify which and how many codons should be removed.
>
> Having reviewed and considered the matter, the court finds the defendants (Monsanto. Co.) have offered clear and convincing evidence which establishes that the specifications of the '600 and '862 patents are not enabling. The court will, therefore, grant the defendants motions and enter an order directing the clerk to enter judgment in favor of defendants and against plaintiff on defendants' counterclaims that the claims of the patent are invalid for lack of enablement.

Thereafter, having the issue whether the claims of Adang's U.S. Patent 5,380,831 are supported by an enabling disclosure before it on appeal from a summary judgment based on the

Delaware district court's decision that the narrower claims of the '600 and '861 patent are not enabled by their supporting specifications, the Federal Circuit remanded the issue to the U.S. District Court for the Southern District of California, stating in Mycogen Plant Sci., Inc. v. Monsanto Co., 252 F.3d 1306, 1317, 58 USPQ2d 1891, 1899 (Fed. Cir. 2001):

> Monsanto's argument on appeal is that the claims are too broad to be enabled by a specification that provides only one example of an embodiment of the invention. The specification of the '831 patent, however, includes more than just one example: it contains codon usage tables, recommendations, on the preferred level of homology, and means for calculating deviation of the frequency of preferred codon usage. The proper resolution of the enablement issue is thus not sufficiently for us to direct the district court to enter summary judgment on that issue before the district court has addressed it. Accordingly, we leave it to the district court to determine whether there is a genuine issue of material fact as to enablement.

The issue here is whether Fischhoff has sustained its burden to show that the specification of Adang's U.S. Patent 5,380,831 would not have enabled persons skilled in the art to make and use the full scope of the subject matter defined by its Claims 1-12 without undue experimentation. To decide the issue, we must consider all evidence of record indicating whether or not modification of a portion of a coding sequence of a gene derived from Bt which encodes an insecticidal protein toxin to yield a modified sequence which (1) contains a greater number of codons preferred by the intended plant host than did the portion of a

coding sequence of a native gene derived from Bt which encodes an insecticidal protein toxin, and/or (2) has a frequency of codon usage which more closely resembles the frequency of codon usage of the plant in which it is to be expressed, necessarily designs a synthetic Bt gene encoding an insecticidal protein toxin which is more highly expressed in plants than a native Bt gene encoding an insecticidal protein toxin. It is anticipated that the evidence and arguments supporting the parties' cases for priority of invention will shed more light on the issue. Accordingly, Fischhoff's Preliminary Motion 5 (Paper No. 82) is DEFERRED TO FINAL HEARING.

N. Fischhoff's Preliminary Motion 9 (Paper No. 87)

By Fischhoff's Preliminary Motion 9 (Paper No. 87), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 are unpatentable under 35 U.S.C. § 112, first paragraph, for noncompliance with its written description requirement. Fischhoff's arguments in support of this motion are best represented by the following quotations:

> Claim 1 of the '831 patent recites that the modified gene sequence "contains a greater number of codons preferred by the intended plant host than did said coding sequence [subjected to analysis]." The '482 application specification does not contain any reference to or definition of the term "greater number of codons preferred by the intended plant host." In the absence of any definition of this

> term, the ordinary meaning of the term "greater number of codons" would include any number of codons including only one codon.

(Paper No. 87, p. 12, second full para.);

> Claim 11 of the '831 patent contains mention of "frequency of codon usage" with no mention of "preferred" as a characteristic of codon usage. The '482 application specification does not contain any reference to or definition of the term "frequency of codon usage."

(Paper No. 87, p. 13, first full para.);

> It was not until the amendment dated June 16, 1993 that applicants Adang et al. added claim 58, among others, which recited "a greater number of codons preferred by the intended plant host than did said coding sequence." New claim 58 was substantially identical to claim 1 of the '831 patent. The amendment cites no specification support for these new claims, but it asserts the following:
>
> > As was discussed during the May 28 interview, simply replacing unfavored codons of the native Bacillus thuringiensis (B.t.) gene with more favored codons of the intended plant host can often effect a number of the other modifications suggested by applicants in the specification.

(Paper No. 87, p. 13, last para.; footnote omitted);

> Adang et al. admitted to what their invention actually was in an amendment dated June 30, 1994, during prosecution of the '191 application [(Adang's U.S. Application 08/057,191)]:
>
> > Applicants discovered that a range of sequence characteristics that are undesirable for expression in plant hosts could be minimized or eliminated by modifying the B.t. gene sequence to reflect a frequency of codon usage more closely resembling that of the intended plant host. [Emphasis added.]

> Claims 1-10, in contrast, read on making just one codon substitution. Claim 12 is even broader in one respect because it reads on making just one nucleotide substitution and claim 11 is still broader because it reads on modifications without any nucleotide substitutions at all.

(Paper No. 87, p. 17; footnote omitted);

> The specification must reasonably convey to one skilled in the relevant art that Adang et al. had possession of a method for enhancing expression of a Bt gene by making a single nucleotide or codon change at the time its application was filed.
>
> Indeed in the office action dated December 11, 1990, in the '482 application, the examiner noted that claims directed to any single factor which was modified were not enabled, because modifying any single factor was unlikely to produce a useful construct. The various factors disclosed by the party Adang et al. as modifications which would produce a more highly expressed gene included the elimination of CUUGG hairpins, CG and TA doublet avoidance, the elimination of polyadenylation sequences, the elimination of polymerase II termination sequences, and the elimination of plant consensus splice sites. The party Adang et al. not only failed to provide any data supporting the operability of any such single modified sequence, but it also failed to note written description support in the specification for making only one type of modification. Clearly, Adang et al. could not point to support in the specification for a single type of modification comprising a change of only one nucleotide.

(Paper No. 87, pp. 21-22; footnote omitted); and

> . . . [I]n response to the same type of undue breadth rejection, the party Adang et al. submitted a declaration by Michael G. Murray, Ph.D., which stated:
>
> > The most important consideration is to more closely approximate the frequency of codon usage of the intended plant host while avoiding sequences known to be undesirable

> to plants. It is obvious that there would be a great number of possibilities for modifying a native B.t. sequence while accomplishing these goals. However, so long as these goals are accomplished, the fact that the resulting synthetic sequences differ would not cause the ordinary skilled artisan to doubt that each synthetic gene would work as taught by the Adang application.
>
> While it is true that there would be a great number of possibilities for modifying a native B.t. gene according to the general teachings in the '831 patent, Dr. Murray's argument implies that multiple changes are needed to "accomplish those goals." In the absence of evidence that a small or specific region in the gene is particularly important to modify, one would not expect that just one or even several modifications would accomplish these goals.

(Paper No. 87, pp. 22-23; footnote omitted).

Accordingly, Fischhoff argues (Paper No. 87, p. 24):

> [T]he specification of the '482 application does not reasonably convey to one skilled in the relevant art that the applicants Adang et al. had possession of a method for modifying Bt sequences comprising a single nucleotide or codon change, at the time that the '482 application was filed. Because they did not have support for the recitation of "greater number of codons" (claim 1) and "wherein the modification step comprises the substitution of at least one nucleotide," (claim 12) each of claims 1-10 and 12 is invalid under the first paragraph of 36 USC 112.

Fischhoff has not satisfied its burden to show that it is entitled to the relief sought. Fischhoff's motion is denied for the reasons stated below.

Fischhoff criticizes Adang's written description of the invention claimed in terms more pertinent to determinations that claims are unpatentable because the full scope of the subject

matter claimed therein would not have been enabled by the supporting specification. Note that Fischhoff argues that "Claim 12 is even broader . . . because it reads on making just one nucleotide substitution and claim 11 is still broader because it reads on modifications without any nucleotide substitutions at all" (Paper No. 87, p. 17); "the examiner noted that claims directed to any single factor which was modified were not enabled, because modifying any single factor was unlikely to produce a useful construct" (Paper No. 87, p. 21); "Adang et al. not only failed to provide any data supporting the operability of any such single modified sequence, but it also failed to note written description support in the specification for making only one type of modification. Clearly, Adang et al. could not point to support in the specification for a single type of modification comprising a change of only one nucleotide" (Paper No. 87, pp. 21-22); and, "In the absence of evidence that a small or specific region in the gene is particularly important to modify, one would not expect that just one or even several modifications would accomplish these goals" (Paper No. 87, p. 23).

On consideration of a movant's case for unpatentability under 35 U.S.C. § 112, first paragraph, its written description requirement and its enablement requirement should not be

confused. The court stated in Vas-Cath, Inc. v. Mahurkar, 935 F.2d 1555, 1563-1564, 19 USPQ2d 1111, 1116-1117 (Fed. Cir. 1991):

> This court in [In re] Wilder[, 736 F.2d 1516, 1520, 222 USPQ 369, 372 (Fed. Cir. 1984), cert. denied, 469 U.S. 1209 (1985),] (and the CCPA before it) clearly recognized, and we hereby reaffirm, that 35 USC 112, first paragraph, requires a "written description of the invention" which is separate and distinct from the enablement requirement. The purpose of the "written description" requirement is broader than to merely explain how to "make and use"; the applicant must also convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession of the invention. The invention is, for purposes of the "written description" inquiry, whatever is now claimed.

Thus, a specification may convey that it was in possession of a broadly claimed invention and satisfy the written description requirement of 35 U.S.C. § 112, first paragraph, without describing all species the broadly claimed invention encompasses. Utter v. Hiraga, 845 F.2d 993, 998, 6 USPQ2d 1709, 1714 (Fed. Cir. 1988). "[T]hat a claim may be broader than the specific embodiment disclosed in a specification is in itself of no moment." In re Rasmussen, 650 F.2d 1212, 1215, 211 USPQ 323, 326 (CCPA 1981). The description requirement of the first paragraph of 35 U.S.C. § 112 is satisfied if one skilled in the art would have understood from the specification as a whole that the broader claim language defines an invention disclosed. Id. at 1216, 211 USPQ at 327, citing In re Smythe, 480 F.2d 1376,

1384, 178 USPQ 279, 285 (CCPA 1973). How close the written description must be to the language of the claimed subject matter to comply with § 112 is determined on a case-by-case basis. In re Smith, 458 F.2d 1389, 1395, 173 USPQ 679, 683 (CCPA 1972). Each case must be decided on its own facts. In re Driscoll, 562 F.2d 1245, 1250, 195 USPQ 434, 438 (CCPA 1977).

In this case, Adang's patent specification teaches:

> The invention disclosed herein comprises a chemically synthesized gene encoding an insecticidal protein which is functionally equivalent to a native insecticidal protein of Bt. This synthetic gene is designed to be expressed in plants at a level higher than a native Bt gene. . . . Preferably, the synthetic gene is at least approximately 85% homologous to an insecticidal protein gene of Bt.

(U.S. Patent 5,380,831, col. 3, l. 56-64);

> In designing synthetic Btt genes of this invention for enhanced expression in plants, the DNA sequence of the native Btt structural gene is modified in order to contain codons preferred by highly expressed plant genes, to attain an A+T content in nucleotide base composition substantially that found in plants, and also preferably to form a plant initiation sequence, and to eliminate sequences that cause destabilization, inappropriate polyadenylation, degradation and termination of RNA and to avoid sequences that constitute secondary structure hairpins and RNA splice sites. In the synthetic genes, codons used to specify a given amino acid are selected with regard to the distribution frequency of codon usage employed in highly expressed plant genes to specify that amino acid. As is appreciated by those skilled in the art, the distribution frequency of codon usage utilized in the synthetic gene is a determinant of the level of expression. Hence, the synthetic gene is designed such that its distribution frequency of codon usage deviates,

> preferably, no more than 25% from that of highly expressed plant genes and, more preferably, no more than about 10%.

(U.S. Patent 5,380,831, col. 4, l. 3-25; emphasis added);

> A structural gene may contain one or more modifications in either the coding or the untranslated regions which could affect the biological activity or the chemical structure of the expression product, the rate of expression or the manner of expression control. Such modifications include, but are not limited to, mutations, insertions, deletions and substitutions of one or more nucleotides.

(U.S. Patent 5,380,831, col. 5, l. 66, to col. 6, l. 5; emphasis added);

> The synthetic Bt genes of the present invention are not considered to be functionally equivalent to native Bt genes, since they are expressible at a higher level in plants than native Bt genes.

(U.S. Patent 5,380,831, col. 6, l. 64-68);

> When synthesizing a gene for improved expression in a host cell it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

(U.S. Patent 5,380,831, col. 7, l. 25-28; emphasis added);

> The term designed to be highly expressed as used herein refers to a level of expression of a designed gene wherein the amount of its specific mRNA transcripts produced is sufficient to be quantified in Northern blots and, thus, represents a level of specific mRNA expressed corresponding to greater than or equal to approximately 0.001% of the poly(A)+ mRNA. To date, natural Bt genes are transcribed at a level wherein the amount of specific mRNA produced is insufficient to be estimated using the Northern blot technique.

(U.S. Patent 5,380,831, col. 8, l. 8-17);

> Thus, this invention is based on the recognition that expression levels of desired, recombinant insecticidal protein in transgenic plants can be improved via increased expression of stabilized mRNA transcripts; and that, conversely, detection of these stabilized RNA transcripts may be utilized to measure expression of translational product (protein).

(U.S. Patent 5,380,831, col. 8, l. 54-60); and

> Experimental evidence obtained from point mutations and deletion analysis has indicated that in eukaryotic genes specific sequences are associated with post-transcriptional processing, RNA destabilization, translational termination, intron splicing and the like. These are preferably employed in the synthetic genes of this invention. In designing a bacterial gene for expression in plants, sequences which interfere with the efficacy of gene expression are eliminated.

(U.S. Patent 5,380,831, col. 10, l. 11-19; emphasis added).

While the above-quoted portions from Adang's U.S. Patent 5,380,831 indicate that point mutations in the native Bt coding sequence for insecticidal protein and/or a synthetic modification of the Bt coding sequence for insecticidal protein wherein the modification comprising the substitution and/or deletion of one nucleotide in the native Bt coding sequence may design a synthetic Bt gene to be more highly expressed in a plant, Fischhoff questions whether Adang's disclosure would have led persons skilled in the art to understand that Adang's invention comprised a method of designing a native Bt coding sequence to be more highly expressed in a plant comprising a point mutation and/or substitution or deletion of one nucleotide. A finding

that persons skilled in the art would have believed that, at the time its patented application was filed, Adang's invention comprised a method of designing native Bt coding sequence to be more highly expressed in a plant comprising a point mutation and/or substitution or deletion of one nucleotide, is supported by the following passages (Adang's U.S. Patent 5,380,831, col. 12, l. 43, to col. 12, l. 21; emphasis added):

> In designing a synthetic gene for expression in plants, attempts are also made to eliminate sequences which interfere with the efficacy of gene expression. Sequences such as the plant polyadenylation signals, . . . polymerase II termination sequence, . . . UCUUCGG hairpins and plant consensus splice sites are highlighted and, if present in the native Btt coding sequence, are modified to eliminate potentially deleterious sequences.
>
> Modifications in nucleotide sequence of the Btt coding region are also preferably made to reduce the A+T content in DNA base composition. . . . Since A+T-rich regions typify plant intergenic regions and plant regulatory regions, it is deemed prudent to reduce the A+T content. . . . .
>
> Also, a single modification (to introduce guanine in lieu of adenine) at the fourth nucleotide position in the Btt coding sequence is made in the preferred embodiment to form a sequence consonant with that believed to function as a plant initiation sequence (Taylor et al. (1987) Mol. Gen. Genet. 210: 572-577) in optimization of expression. . . .
>
> Not all of the above-mentioned modifications of the natural Bt gene must be made in constructing a synthetic Bt gene in order to obtain enhanced expression. For example, a synthetic gene may be synthesized for other purposes in addition to that of achieving enhanced levels of expression. Under these conditions, the original sequence of the

> natural Bt gene may be preserved within a region of DNA corresponding to one or more, but not all, segments used to construct the synthetic gene. Depending on the desired purpose of the gene, modification may encompass substitution of one or more, but not all, of the oligonucleotide segments used to construct the synthetic gene by a corresponding region of natural Bt sequence.

Adang's patent specifies (U.S. Patent 5,380,831, col. 15, l. 30-55):

> This invention combines the specific teachings of the present disclosure with a variety of techniques and expedients known in the art. The choice of expedients depends on variables such as the choice of insecticidal protein from a Bt strain, the extent of modification in preferred codon usage, manipulation of sequences considered to be destabilizing to RNA or sequences prematurely terminating transcription, insertions of restriction sites within the design of the synthetic gene to allow future nucleotide modifications, additions of introns or enhancer sequences to the 5' and/or 3' ends of the synthetic gene, the promoter region, the host in which a promoter region/structural gene combination is expressed, and the like. . . . The fundamental aspect of the present invention is the ability to synthesize a novel gene coding for an insecticidal protein, designed so that the protein will be expressed at an enhanced level in plants, yet so that it will retain its inherent property of insect toxicity and retain or increase its specific insecticidal activity.

Persons skilled in the art would have understood from the teaching of Adang's patent specification as a whole that the invention described therein minimally comprises a single modification of a native Bt gene encoding a insecticidal protein, e.g., the introduction, deletion or substitution at a single

nucleotide position, to enhance expression of the insecticidal protein in a plant by a detectable amount of mRNA transcripts.

The more debatable issue is, as Fischhoff suggests, whether Adang's patent specification would have enabled persons skilled in the art to make and use the full scope of the invention described and claimed. That issue is deferred to final hearing. Fischhoff has not otherwise shown that Adang's patent specification does not satisfy the written description of 35 U.S.C. § 112, first paragraph, for the subject matter claimed. Accordingly, Fischhoff's Preliminary Motion 9 (Paper No. 87) is DENIED.

O. Fischhoff's Preliminary Motion 7 (Paper No. 85)

By Fischhoff's Preliminary Motion 7 (Paper No. 85), Fischhoff moves under 37 CFR § 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 (FX 11), issued January 10, 1995, are unpatentable under 35 U.S.C. § 102 or § 103 (Paper No. 85, p. 1). Consideration of this motion is deferred to final hearing.

When Fischhoff filed its motion (Paper No. 85), Adang opposed the motion (Paper No. 68), and Fischhoff replied (Paper No. 107), neither party to this interference appears to have had the benefit of the interpretations of terms in its claims by the U.S. District Courts for the District of Delaware and the

Southern District of California and the decisions of the U.S. Court of Appeals for the Federal Circuit in review thereof. It is improper to consider the patentability of claimed subject matter under the first paragraph of 35 U.S.C. § 112, 35 U.S.C. § 102, and 35 U.S.C. § 103 before the scope and content of the subject matter claimed is understood. See Panduit Corp. v. Dennison Mfg., 810 F.2d 1561, 1567-1568, 1 USPQ2d 1593, 1597 (Fed. Cir.), cert. denied, 481 U.S. 1052 (1987)(footnote omitted):

> Analysis begins with a key legal question - what is the invention claimed? Courts are required to view the claimed invention as a whole. 35 U.S.C. § 103. Claim interpretation, in light of the specification, claim language, other claims, and prosecution history, is a matter of law, and will normally control the remainder of the decisional process.

See also In re Geerdes, 491 F.2d 1260, 1262, 180 USPQ 789, 791 (CCPA 1974)("[b]efore considering the rejections under 35 U.S.C. 103 and 112, we must first decide [what] . . . the claims include in their scope"); In re Moore, 439 F.2d 1232, 1235, 169 USPQ 236, 238 (CCPA 1971)("the claims must be analyzed first in order to determine exactly what subject matter they encompass"); and In re Steele, 305 F.2d 859, 862, 134 USPQ 292, 295 (CCPA 1962):

> We do not think a rejection under 35 U.S.C. 103 should be based on . . . speculations and assumptions. We think [it] . . . wrong in relying on what at best are speculative assumptions as to the meaning of the claims and basing a rejection under 35 U.S.C. 103 thereon.

Along with filing this motion, Fischhoff moved under 37 CFR § 1.633(a) for judgment that the Adang's claims designated as corresponding to the count do not distinctly claim and particularly define the subject matter Adang regards as its invention (Fischhoff's Preliminary Motion 3 (Paper No. 80)). Therein, Fischhoff argued that the subject matter defined by Adang's claims is vague and indefinite. Considering that view, it must be assumed that the respective positions of the parties supporting and opposing this motion are based on speculation. Accordingly, it is most appropriate, informative, and efficient at this time to defer consideration of the issues presented by this motion until after the parties have had ample opportunities to comprehend the Federal Circuit's review of district courts' decisions interpreting language common to, and the scope and content of, the parties' claims designated as corresponding to the count, and reconsider and support their respective positions in that light.

> Another key preliminary legal inquiry is - what is the prior art? ["Under 35 U.S.C. § 103, the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of skill in the pertinent art resolved. Against this background, the obviousness or nonobviousness of the subject matter is determined." Graham v. John Deere Co., 383 U.S. 1, 17-18 (1966).] Before answering Graham's "content" inquiry, it must be known whether a patent or publication is in the prior art under 35 U.S.C. § 102, - a legal question.

Panduit Corp. v. Dennison Mfg., 810 F.2d at 1568, 1 USPQ2d at 1597. The primary purpose of an interference proceeding is to determine priority of a commonly claimed invention, i.e., what constitutes prior art under 35 U.S.C. § 102(g). Accordingly, the issue whether Adang's claims are unpatentable under 35 U.S.C. § 103 cannot and should not be decided until the full scope and content of the prior art have been determined; the differences between the prior art and the claims at issue have been ascertained; and the level of skill in the pertinent art is resolved. Accordingly, Fischhoff' Preliminary Motion 7 (Paper No. 85) is DEFERRED TO FINAL HEARING.

P. Fischhoff's 37 CFR § 1.641(a) Request (Paper No. 110)

By Fischhoff's 37 CFR § 1.641(a) Request (Paper No. 110), Fischhoff asks the APJ "to set a time period within which each party may develop testimony and present arguments and evidence on the apparent invalidity of" Claims 1-12 of Adang's U.S. Patent 5,380,831 under 35 U.S.C. § 112, first paragraph, for noncompliance with its best mode requirement (Paper No. 110, pp. 6-7, bridging para.). As support therefor, Fischhoff argues that the APJ, exercising discretion which is permissible under 37 CFR § 1.641(a), should "conclude" that Claims 1-12 of Adang's U.S. Patent 5,380,831 are invalid for noncompliance with the best mode requirement of 35 U.S.C. § 112, first paragraph.

According to Fischhoff, it is "apparent" from the evidence that Adang's patent specification does not satisfy the best mode requirement of 35 U.S.C. § 112, first paragraph (Paper No. 110, pp. 6-7, bridging para.). Nevertheless, given that Fischhoff itself has not moved under 37 CFR 1.633(a) for judgment that Claims 1-12 of Adang's U.S. Patent 5,380,831 are unpatentable under 35 U.S.C. § 112, first paragraph, it appears that (1) the evidence to which Fischhoff points may not be sufficient to establish the invalidity of Claims 1-12 of Adang's U.S. Patent 5,380,831 under 35 U.S.C. § 112, first paragraph (best mode requirement), and (2) Fischhoff itself does not consider the evidence sufficient to establish invalidity under 35 U.S.C. § 112, first paragraph. Indeed, Fischhoff's request is confusing.

In Eli Lilly & Co. v. Barr Labs., 251 F.3d 955, 58 USPQ2d 1869 (Fed. Cir. 2001), the court said at 963, 58 USPQ2d at 1874:

> Our case law explicating the best mode requirement focuses on a two-prong inquiry. . . . First, the factfinder must determine whether, at the time of filing the application, the inventor possessed a best mode for practicing the invention. . . . Second, if the inventor possessed a best mode the factfinder must determine whether the written description disclosed the best mode such that one reasonably skilled in the art could practice it. . . . The first prong involves a subjective inquiry, focusing on the inventor's state of mind at the time of filing. . . . The second prong involves an objective inquiry, focusing on the scope of the claimed invention and the level of skill in the art. . . .

> With respect to the second prong of the best mode requirement, the extent of information that an inventor must disclose depends on the scope of the claimed invention. . . . .

In this case, it appears that the synthetic genes comprising DNA sequences of Figure 1 of Adang's U.S. Patent 5,380,831 which span nucleotides 1 through 1793 and 1 through 1833 represent specific examples of useful products which can be made by Adang's generally claimed method of designing a synthetic Bt gene to be more highly expressed in plants. However, Adang's specification makes clear that useful products made in accordance with the claimed invention are much broader in scope than the patent specification's examples, and the synthetic Bt genes comprising the DNA sequences depicted in Figure 1 and defined by Claims 13 and 14 of Adang's patent do not necessarily represent the best mode of designing a synthetic Bt gene to be more highly expressed either in plants in general or in any particular kind of plant. Adang's patent specification teaches (Adang's U.S. Patent 5,380,831, col. 15, l. 30-55):

> This invention combines the specific teachings of the present disclosure with a variety of techniques and expedients known in the art. The choice of expedients depends on variables such as the choice of insecticidal protein from a Bt strain, the extent of modification in preferred codon usage, manipulation of sequences considered to be destabilizing to RNA or sequences prematurely terminating transcription, insertions of restriction sites within the design of the synthetic gene to allow future nucleotide modifications, addition of introns or enhancer sequences to the 5' and/or

> 3' ends of the synthetic structural gene, the promoter region, the host in which a promoter region/structural gene combination is expressed, and the like. As novel insecticidal proteins and toxic polypeptides are discovered, and as sequences responsible for enhanced cross-expression (expression of a foreign structural gene in a given host) are elucidated, those of ordinary skill will be able to select among those elements to produce "improved" synthetic genes for desired proteins having agronomic value. The fundamental aspect of the present invention is the ability to synthesize a novel gene coding for an insecticidal protein, designed so that the protein will be expressed at an enhanced level in plants, yet so that it will retain its inherent property of insect toxicity and retain or increase its specific insecticidal activity.

From the above quotation from Adang's patent, it is not at all apparent that Adang possessed a best mode generally for practicing the claimed invention at the time its patent application was filed. It is even less apparent that the invention Adang claims can have a best mode without specifying the type of plant in which the synthetic Bt gene is to be highly expressed, the particular extent and/or kind of insect infestation, the particular toxin the Bt gene is to encode, the particular amount and/or strength of the toxin the plant must produce relative to the kind of infestation and plant, etc.. Given the number of expedients influencing the design of the synthetic Bt gene to be produced by the methods generally claimed in Adang's patent, it is not apparent that Adang possessed a best mode of practicing the method generally claimed.

Moreover, in its broadest aspects, the methods claimed in Adang's patent specify that the native Bt gene is to be modified to yield a modified sequence which contains a greater number of codons preferred by the intended plant host" (Claim 1 of Adang's U. S. Patent 5,380,831) or to have a frequency of codon usage more closely resembling the frequency of codon usage of the plant in which it is to be expressed (Claim 11 of Adang's U. S. Patent 5,380,831). It is not apparent that the examples in Adang's patent do not set forth the best mode of carrying out the subject matter Adang most broadly claims to be its invention.

Since it is not apparent from the evidence to which Fischhoff points that Claims 1-12 of Adang's patent are unpatentable under 35 U.S.C. § 112, first paragraph (best mode requirement), there is no reason for the APJ to exercise discretion permissible under 37 CFR § 1.641(a) and "conclude" that Claims 1-12 of Adang's U.S. Patent 5,380,831 are invalid for noncompliance with the best mode requirement of 35 U.S.C. § 112, first paragraph, or "to set a time period within which each party may develop testimony and present arguments and evidence . . . [relative to] the apparent invalidity of" Claims 1-12 of Adang's U.S. Patent 5,380,831 under 35 U.S.C. § 112, first paragraph, for noncompliance with its best mode requirement. Fischhoff's 37 CFR § 1.641(a) Request (Paper No. 110) is DENIED.

Q. Fischhoff's Motion Under 37 CFR § 1.635 For Temporary Stay Of Interference Under 37 CFR § 1.645(d) (Paper No. 118)

Fischhoff moves for a temporary stay of this interference proceeding pending the decision of the U.S. District Court for the Southern District of California in Mycogen Plant Sci., Inc., and Agrigenetics, Inc. v. Monsanto Co., Case No. 95-0653j, on Monsanto's motion for summary judgment that all claims of Adang's U.S. Patent 5,380,831 are invalid under 35 U.S.C. §§ 102(g) and 103 because of prior invention thereof by Fischhoff. In that case, Mycogen Plant Science Inc. and Agrigenetics Inc. sued Monsanto Company for infringement of its patent (Adang et al., U.S. Patent 5,380,831, issued January 10, 1995, from U.S. Application 08/057,191, filed May 3, 1993). The U.S. District Court for the Southern District of California entered an order (Mycogen Plant Sci., Inc. v. Monsanto Co., No. 95-CV-653 (S.D. Cal. Nov. 10, 1999)(Paper No. 127, Exh. A)) granting defendant's motion for summary judgment that Claims 1-12 of Mycogen's '831 patent are invalid under 35 U.S.C. § 102(g) and/or § 103 because Monsanto invented the subject matter thereof before Mycogen, as determined by the U.S. District Court for the District of Delaware in Mycogen Plant Sci., Inc. v. Monsanto Co., 61 F. Supp. 2d 199 (D. Del. 1999), which was affirmed in Mycogen Plant Sci., Inc. v. Monsanto Inc., 243 F.3d 1316, 58 USPQ2d 1030 (Fed. Cir.

2001), and denied defendant's motion for summary judgment that the contested claims of Mycogen's '831 patent are invalid for noncompliance with the enablement requirement of the first paragraph of 35 U.S.C. § 112 as moot (Paper No. 127, Exh. A).

On appeal from the decision of the U.S. District Court for the Southern District of California on motion for summary judgment in Mycogen Plant Sci., Inc. v. Monsanto Co., No. 95-CV-653 (S.D. Cal. Nov. 10, 1999)(Paper No. 127, Exh. A), the U.S. Court of Appeals for the Federal Circuit affirmed-in-part, reversed-in-part, and remanded. Mycogen Plant Sci., Inc. v. Monsanto Co., 252 F.3d 1306, 1309, 58 USPQ2d 1891, 1892-1893 (Fed. Cir. 2001). The Federal Circuit concluded at 1309, 58 USPQ2d at 1893, that:

> . . . the district court improperly resolved disputed questions of material fact pertaining to the issue of prior invention, and we therefore reverse the court's ruling on summary judgment that the '831 patent is invalid under 35 U.S.C. § 102(g). We decline to affirm the summary judgment of invalidity on the alternative ground of non-enablement, as urged by Monsanto, but leave to the district court the task of determining in the first instance whether there is a genuine issue of material fact as to enablement based on its assessment of the evidence presented to it in the summary judgment proceeding.

Id. at 1310, 58 USPQ2d at 1894, the Federal Circuit explained:

> We agree with the district court that collateral estoppel requires the court to conclude that Monsanto reduced the invention [claimed in the Mycogen's '831 patent] to practice before Mycogen, and that collateral estoppel does not resolve the question whether Mycogen

> was the first to conceive and then was diligent during the critical period. On the merits of the summary judgment question, however, we do not agree that Monsanto has met its burden of showing that there are no issues of material fact regarding whether Mycogen was the first to conceive the invention and then diligently reduce it to practice.

Because Fischhoff moves for a temporary stay of this interference proceeding pending the decision of the U.S. District Court for the Southern District of California in Mycogen Plant Sci., Inc., and Agrigenetics, Inc. v. Monsanto Co., Case No. 95-0653j, on Monsanto's motion for summary judgment that all claims of Adang's U.S. Patent 5,380,831 are invalid under 35 U.S.C. §§ 102(g) and 103 because of prior invention thereof by Fischhoff, the decision of the U.S. District Court for the Southern District of California on Monsanto's motion for summary judgment that all claims of Adang's U.S. Patent 5,380,831 are invalid under 35 U.S.C. §§ 102(g) and 103 was entered November 10, 1999, and the district court's ruling on summary judgment that the '831 patent is invalid under 35 U.S.C. § 102(g)/103 was reversed on appeal to the U.S. Court of Appeals for the Federal Circuit in Mycogen Plant Sci., Inc. v. Monsanto Co., 252 F.3d 1306, 1309, 58 USPQ2d 1891, 1892-1893 (Fed. Cir. 2001), Fischhoff's motion for a temporary stay of this interference proceeding pending the decision of the U.S. District Court for the Southern District of California in Mycogen Plant

Sci., Inc., and Agrigenetics, Inc. v. Monsanto Co., Case No. 95-0653j, on Monsanto's motion for summary judgment that all claims of Adang's U.S. Patent 5,380,831 are invalid under 35 U.S.C. §§ 102(g) and 103 because of prior invention thereof by Fischhoff (Paper No. 118) is DISMISSED.

R. Fischhoff's Motion Under 37 CFR § 1.635 For Temporary Stay Of Interference Under 37 CFR § 1.645(d) Pending Outcome Of Federal Circuit Appeal Of The District Court Judgment Of Invalidity Of Claims Of The Involved Adang Patent Based On Prior Invention By Fischhoff (Paper No. 127)

The decision of the U.S. District Court for the Southern District of California on Monsanto's motion for summary judgment that all claims of Adang's U.S. Patent 5,380,831 are invalid under 35 U.S.C. §§ 102(g) and 103 was entered November 10, 1999, and the district court's ruling on summary judgment that the '831 patent is invalid under 35 U.S.C. § 102(g)/103 was reversed on appeal to the U.S. Court of Appeals for the Federal Circuit in Mycogen Plant Sci., Inc. v. Monsanto Co., 252 F.3d 1306, 58 USPQ2d 1891 (Fed. Cir. 2001). Thus, Fischhoff's motion for a temporary stay is moot. Fischhoff's motion (Paper No. 127) is DISMISSED.

4. Summary of Decisions on Outstanding Requests and Motions

A. Adang Preliminary Motion 1 (Paper No. 45) DENIED;
B. Adang Preliminary Motion 2 (Paper No. 46) DISMISSED;

C. Adang Preliminary Motion 3 (Paper No. 47) DENIED;

D. (1) Fischhoff First Rule 642 Request (Paper No. 78) DISMISSED;
(1) Fischhoff Second Rule 642 Request (Paper No. 79) DISMISSED;

E. Fischhoff Preliminary Motion 4 (Paper No. 81) DENIED;

F. Fischhoff Preliminary Motion 8 (Paper No. 86) DISMISSED;

G. Fischhoff Preliminary Motion 11 (Paper No. 89) DISMISSED;

H. Fischhoff Preliminary Motion 12 (Paper No. 60) DISMISSED;

I. Adang's Miscellaneous Rule 635 Motion (Paper No. 116) GRANTED;

J. Fischhoff Preliminary Motion 6 (Paper No. 83) DENIED;

K. Fischhoff Preliminary Motion 10 (Paper No. 88) DEFERRED TO FINAL HEARING;

L. Fischhoff Preliminary Motion 3 (Paper No. 80) DENIED;

M. Fischhoff Preliminary Motion 5 (Paper No. 82) DEFERRED TO FINAL HEARING;

N. Fischhoff Preliminary Motion 9 (Paper No. 87) DENIED;

O. Fischhoff Preliminary Motion 7 (Paper No. 85) DEFERRED TO FINAL HEARING;

P. Fischhoff Rule 641 Request (Paper No. 110) DENIED;

Q. Fischhoff Miscellaneous Rule 635 Motion (Paper No. 118) DISMISSED;

R. Fischhoff Miscellaneous Rule 635 Motion (Paper No. 118) DISMISSED.

5. Order

This interference has been redeclared as Barton (U.S. Application 07/827,906) or Fischhoff (U.S. Application 08/434,105) v. Adang (U.S. Patent 5,380,831). New Count 2 defines the interfering subject matter:

Count 2

Any one of Claims 1-4, 7, and 15-22 of Barton et al.'s Application 07/827,906, filed January 30, 1992;

- or -

Any one of Claims 3, 5, and 39-43 of Fischhoff et al.'s Application 07/827,906, filed January 30, 1992;

- or -

Any one of Claims 1-14 of Adang et al.'s U.S. Patent 5,380,831, which issued January 10, 1995, from U.S. Application 08/057,191, filed May 3, 1993.

The claims of the parties which have been designated as corresponding to this count are:

Barton: Claims 1-4, 7, and 15-22

Fischhoff: Claims 3, 5, and 39-43

Adang: Claims 1-14.

In addition, this Decision and Order on Preliminary and Miscellaneous Motions and Requests invites the parties to consider the relationship of the subject matter defined by Count 2 of this interference to subject matter claimed in

Mycogen's U.S. Patents 6,013,523 and 6,015,891 and comment thereon.

Accordingly, it is

ORDERED that no later than thirty (30) days from the date of this Decision and Order on Preliminary and Miscellaneous Motions and Requests, the parties to this interference shall specify whether the time for filing preliminary motions should be extended;

FURTHER ORDERED that no later than thirty (30) days from the date of this Decision and Order on Preliminary and Miscellaneous Motions and Requests, the parties to this interference shall specify what additional preliminary motions, if any, and supporting evidence, if any, need be filed in this newly declared interference;

FURTHER ORDERED that no later than thirty (30) days from the date of this Decision and Order on Preliminary and Miscellaneous Motions and Requests, the parties to this interference shall explain why the additional preliminary motions and supporting evidence specified are necessary to, and should be filed in, this interference proceeding, and

FURTHER ORDERED that no later than thirty (30) days from the date of this Decision and Order on Preliminary and Miscellaneous Motions and Requests, the parties to this interference shall

recommend time periods for filing the specified additional preliminary motions, supporting evidence, oppositions, replies, motions to suppress evidence, etc.

Teddy S. Gron
Administrative Patent Judge

Interference 103,781

COUNSEL FOR JUNIOR PARTY BARTON et al.:

Lawrence M. Lavin, Jr., Esq.
Monsanto Agricultural Company
700 Chesterfield Parkway North - BB4F
St. Louis, MO 63198

or

COUNSEL FOR JUNIOR PARTY FISHHOFF et al.:

Lawrence M. Lavin, Jr., Esq.
Monsanto Company
700 Chesterfield Parkway North, BB4F
St. Louis, MO 63198

and

COUNSEL FOR SENIOR PARTY ADANG et al.:

Oliver R. Ashe, Jr., Esq.
Greenblum & Bernstein, P.L.C.
1941 Roland Clarke Place
Reston, VA 20191

APPENDIX A

US006015891A

United States Patent [19]

Adang et al.

[11] Patent Number: 6,015,891

[45] Date of Patent: *Jan. 18, 2000

[54] SYNTHETIC INSECTICIDAL CRYSTAL PROTEIN GENE HAVING A MODIFIED FREQUENCY OF CODON USAGE

[75] Inventors: **Michael J. Adang**, Athens, Ga.; **Elizabeth E. Murray**, Madison, Wis.

[73] Assignee: **Mycogen Plant Science, Inc.**, San Diego, Calif.

[*] Notice: This patent is subject to a terminal disclaimer.

[21] Appl. No.: **08/705,438**

[22] Filed: **Aug. 29, 1996**

Related U.S. Application Data

[60] Division of application No. 08/369,835, Jan. 6, 1995, Pat. No. 5,567,600, which is a continuation-in-part of application No. 08/057,191, May 3, 1993, Pat. No. 5,380,831, which is a continuation of application No. 07/827,844, Jan. 28, 1992, abandoned, which is a continuation of application No. 07/242,482, Sep. 9, 1988, abandoned.

[51] **Int. Cl.**[7] **C12N 15/00**; C12N 15/32; C12N 15/82

[52] **U.S. Cl.** **536/23.71**; 435/440; 435/468

[58] **Field of Search** 536/23.71; 435/172.3, 435/320.1, 69.1, 419, 468, 440; 800/205, 279, 302

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,356,270	10/1982	Itakura	435/317
4,888,282	12/1989	Beremand et al.	435/193
5,254,799	10/1993	De Greve	800/205
5,276,268	1/1994	Strauch et al.	800/205
5,380,831	1/1995	Adang et al.	536/23.71
5,567,600	10/1996	Adang et al.	536/23.71

FOREIGN PATENT DOCUMENTS

0305275 3/1989 European Pat. Off. .

OTHER PUBLICATIONS

Sekar, V. et al. (1987) "Molecular cloning and characterization of the insecticidal crystal protein gene of *Bacillus thuringiensis* var. tenebrionis" Pro. Natl. Acad. Sci. USA 84:7036–7040.

McPherson, S. A. et al. (1988) "Characterization of the Colopteran–specific protein gene of *Bacillus thuringiensis* var. Tenebrionis" Biotechnology 6(1):61–6.

Vaeck, M. et al. (1987) "Transgenic Plants protected from insect attach" Nature 328:33–37.

Murray, E. E. et al. (1989) "Codon usage in plant genes" Nucleic Acids Research 17(2):477–498.

Perlak, F. J. et al. (1991) "Modification of the coding sequence enhances plant expression of insect control protein genes" Proc. Natl. Acad. Sci. USA 88:3324–3328.

Hernastadt et al. (1986) Biotechnology 4:305–308.

Sekar et al. (1987) Proc. Natl. Acad. Sci. 84:7036–7040.

McPherson et al. (1988) Biotechnology 6:61–66.

Vaeck et al. (1987) Nature 328:33–37.

Barton et al. (1987) Plant Physiol. 85: 1103–1109.

Hofte et al. (1988) FEBS Lett. 226:364–370.

Bennetzen and Hall (1982) J. Biol. Chem. 257:3026–3031.

Hoekema et al. (1987) Mol. Cell. Biol. 7:2914–2924.

Boudraa (1987) Genet. Sel. Evol. 19:143–154.

Grantham et al. (1986) Oxford Surveys in Evol. Biol. 3:48–81.

Taylor et al. (1987) Mol. Gen. Genet. 210:572–577.

Joshi (1987) Nucl. Acids Res. 15:9627–9640.

Vankan and Fillipowicz (1988) Embo J. 7:791–799.

Tuerk et al. (1988) Proc. Natl. Acad. Sci. 85:1364–1368.

Krieg et al. (1983) Z. Angew. Entomol. 96:500–508.

Bernahard (1986) FEMS Microbiol. Lett. 33:261–265.

Neill et al. (1987) Gene 55:303–317.

Rothstein et al. (1987) Gene 55:353–356.

Coraggio et al. (1986) EMBO J. 5:459–465.

Fuzakawa et al. (1987) FEBS Lett. 224:125–127.

Vies et al. (1986) EMBO J. 5:2439–2444.

Gatenby et al. (1987) Eur. J. Biochem. 168:227–231.

Viotti et al. (1978) Biochim. Biophys. Acta 517:125–132.

Hames and Higgins (eds) (1985) *Nucleic Acid Hybridisation*, IRL Press, Oxford, UKe.

Caruthers, M. (1983) In Methodology of DNA and RNA Sequencing, Chapter 1, Weissman (ed.), Praeger Publishers, New Yor.I.

Lo et al. (1984) Proc. Natl. Acad. Sci. 81:2285–2289.

Grantham et al. (1985) Bull. inst. Pasteur 83:95–148.

Matsuoka et al. (1987) J. Biochem. 102:673–676.

Mandecki et al. (1985) Proc. Natl. Acad. Sci. 82:3543–3547.

Feretti et al. (1986) Proc. Natl. Acad. Sci. 83:599–603.

Paszkowski, J. et al. (1984) EMBO J. 3: 2717.

Fromm, M. et al. (1985) Proc. Natl. Acad. Sci. USA 82:5824.

Crossway, A. et al. (1986) Mol. Gen. Genet. 202:179.

Hooykaas–Van Slogteren, G. et al. (1984) Nature 311:763.

Dandekar, A. et al. (1987) Biotechnology 5:587.

An, G. et al. (1985) EMBO J. 4:277.

Herrera–Estrella, L. et al. (1983) Nature 303:209E.

Herrera–Estrella, L. et al. (1983) EMBO J. 2:987.

Herrera–Estrella, L. et al. (1985) in Plant Genetic Engineering, Cambridge University Press, New York, p. 63.

asadaban, M. J. and Cohen, S.N. (1980) J. Mol. Biol. 138:179–207.

Marsh, J. L. et al. (1984) Gene 32:481–485.

Kronstad et al. (1983) J. Bacteriol. 154: 419–428.

Adang et al. (1987) in *Biotechnology in Invertebrate Pathology and Cell Culture*, K. Maramorosh (ed.) Academic Press, Inc. New York, pp 85–99.

(List continued on next page.)

Primary Examiner—Lynette R. F. Smith
Assistant Examiner—Amy J. Nelson
Attorney, Agent, or Firm—Saliwanchik, Lloyd & Saliwanchik

[57] **ABSTRACT**

Synthetic *Baccilus thuringiensis* toxin genes designed to be expressed in plants at a level higher than naturally-occurring *Bt* genes are provided. These genes utilize codons preferred in highly expressed monocot or dicot proteins.

6 Claims, 5 Drawing Sheets

OTHER PUBLICATIONS

Schnepf and Whitteley (1985) J. Biol. Chem 260:6273–6280.
Pfitzinger et al. (1987) Nucl. Acids Res. 15: 1377–1386.
Sharp and Li (1986) Nucl. Acids Res. 14:7734–7749.
Matteucci et al. (1981) J. Am. Chem. Soc. 103:3185–3192.
Beaucage et al. (1981) tetrahedron Lett. 22: 1859–1862.
McBride et al. (1988) Biotechniques 6:362–367.
Maniatis (1982) *Cloning Manual*, Fritsch and Sambrook (eds.), Cold Spring Harbor Laboratory, Cold Spring Harebor, New York.
Sanger et al. (1977) proc. Natl. Acad. Sci. 74:5463–5467.
Russell and Bennett (1982) Gene 20: 231–244.
Yanisch–Perron, C. et al. (1985) Gene 33: 103–119.
Barnes (1986) J. Cell Biochem . O (10 part C) 47.
Boswell et al. in *Computational Molecular Biology, Sources and Methods for Sequence Analysis*(Leak, ed.) Oxford University Press, Oxford, 1986, pp. 170–171.
Schnepf HE. "*Bacillus thuringiensis* toxins: Regulation, activities and structural diversity." Current Opinion Biotech. 6: 305–312, 1995.

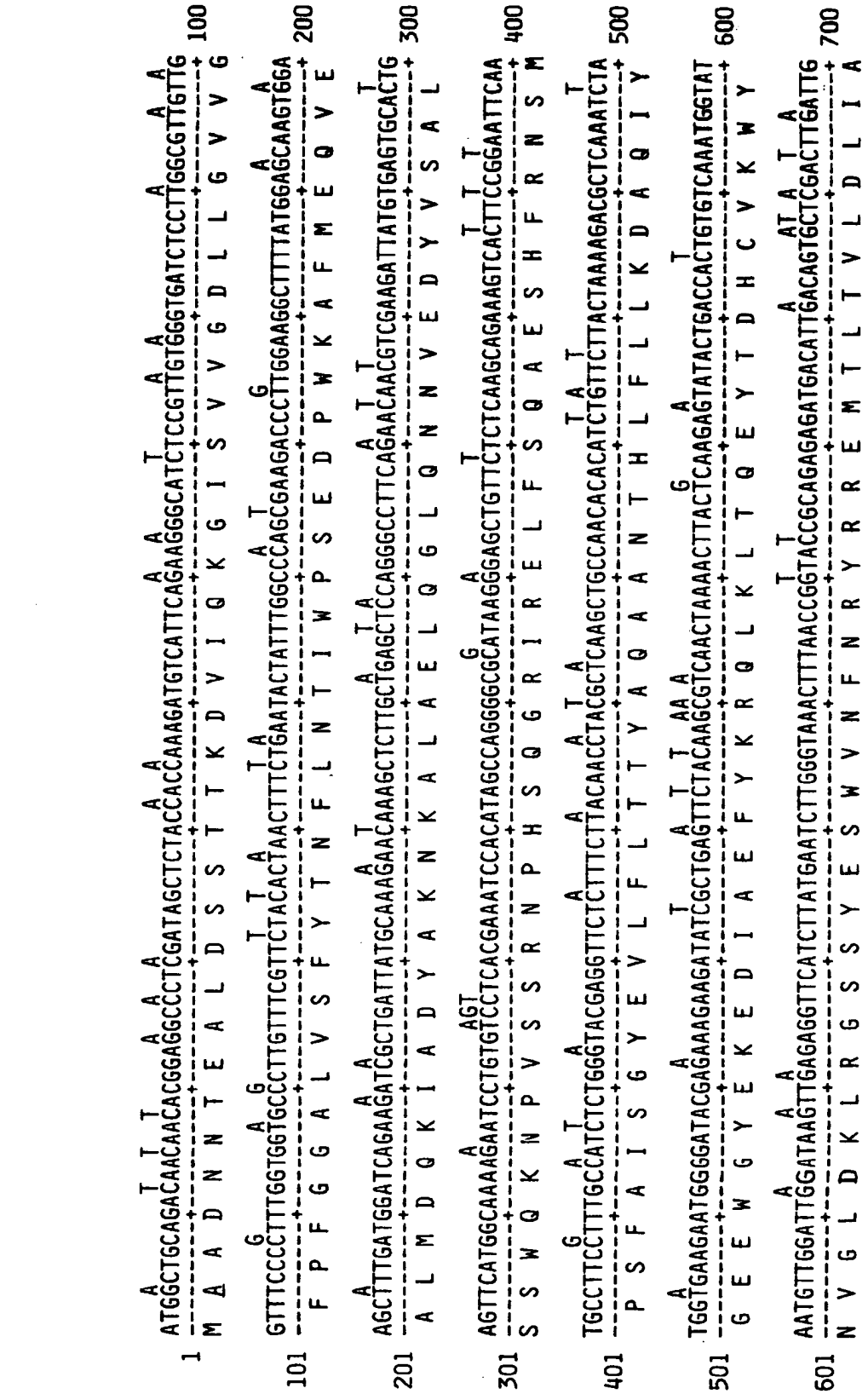


FIG.IA

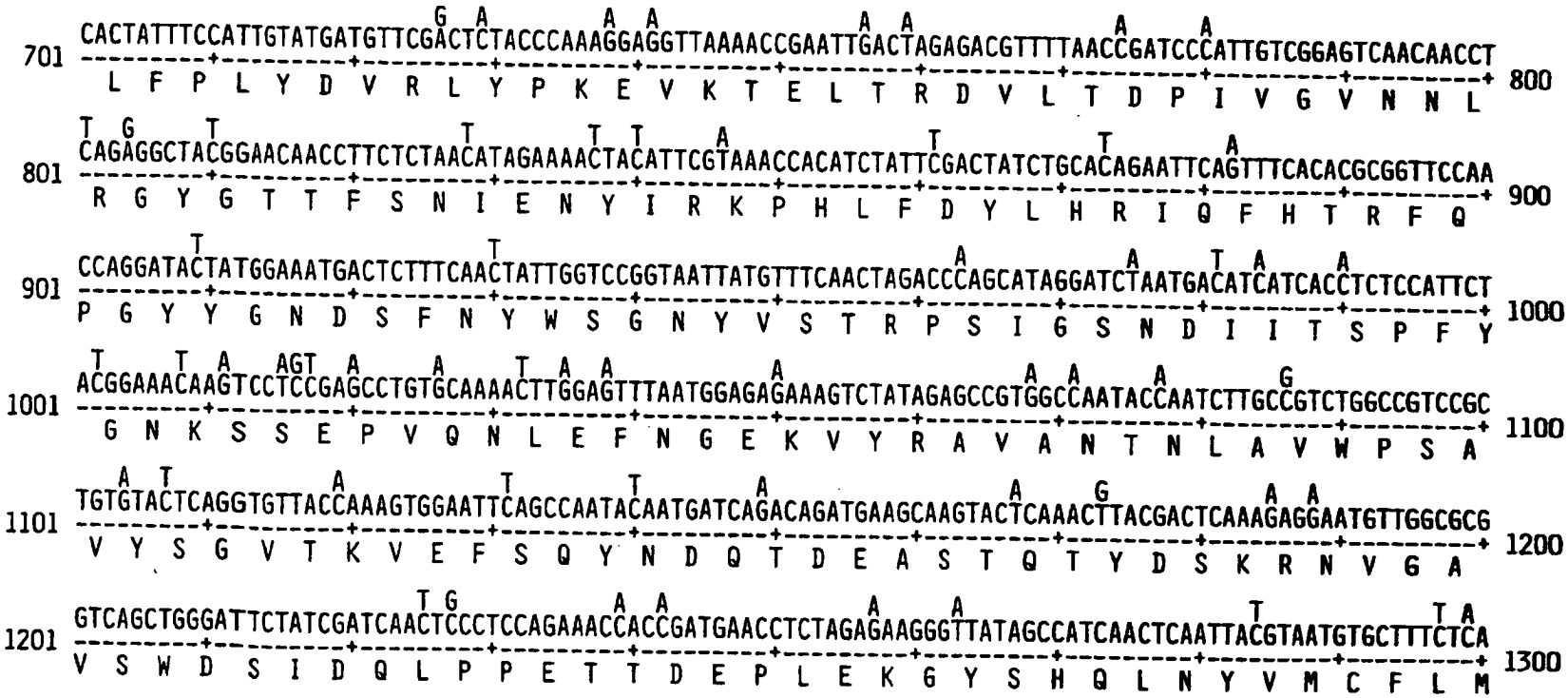

701
CACTATTTCCATTGTATGATGTTCGACTCTACCCAAAGGAGGTTAAAACCGAATTGACTAGAGACGTTTTAACCGATCCCATTGTCGGAGTCAACAACCT
L F P L Y D V R L Y P K E V K T E L T R D V L T D P I V G V N N L
800
801
CAGAGGCTACGGAACAACCTTCTCTAACATAGAAAACTACATTCGTAAACCACATCTATTCGACTATCTGCACAGAATTCAGTTTCACACGCGGTTCCAA
R G Y G T T F S N I E N Y I R K P H L F D Y L H R I Q F H T R F Q
900
901
CCAGGATACTATGGAAATGACTCTTTCAACTATTGGTCCGGTAATTATGTTTCAACTAGACCCAGCATAGGATCTAATGACATCATCACCTCTCCATTCT
P G Y Y G N D S F N Y W S G N Y V S T R P S I G S N D I I T S P F Y
1000
1001
ACGGAAACAAGTCCTCCGAGCCTGTGCAAAACTTGGAGTTTAATGGAGAGAAAGTCTATAGAGCCGTGGCCAATACCAATCTTGCCGTCTGGCCGTCCGC
G N K S S E P V Q N L E F N G E K V Y R A V A N T N L A V W P S A
1100
1101
TGTGTACTCAGGTGTTACCAAAGTGGAATTCAGCCAATACAATGATCAGACAGATGAAGCAAGTACTCAAACTTACGACTCAAAGAGGAATGTTGGCGCG
V Y S G V T K V E F S Q Y N D Q T D E A S T Q T Y D S K R N V G A
1200
1201
GTCAGCTGGGATTCTATCGATCAACTCCCTCCAGAAACCACCGATGAACCTCTAGAGAAGGGTTATAGCCATCAACTCAATTACGTAATGTGCTTTCTCA
V S W D S I D Q L P P E T T D E P L E K G Y S H Q L N Y V M C F L M
1300

FIG. 1B

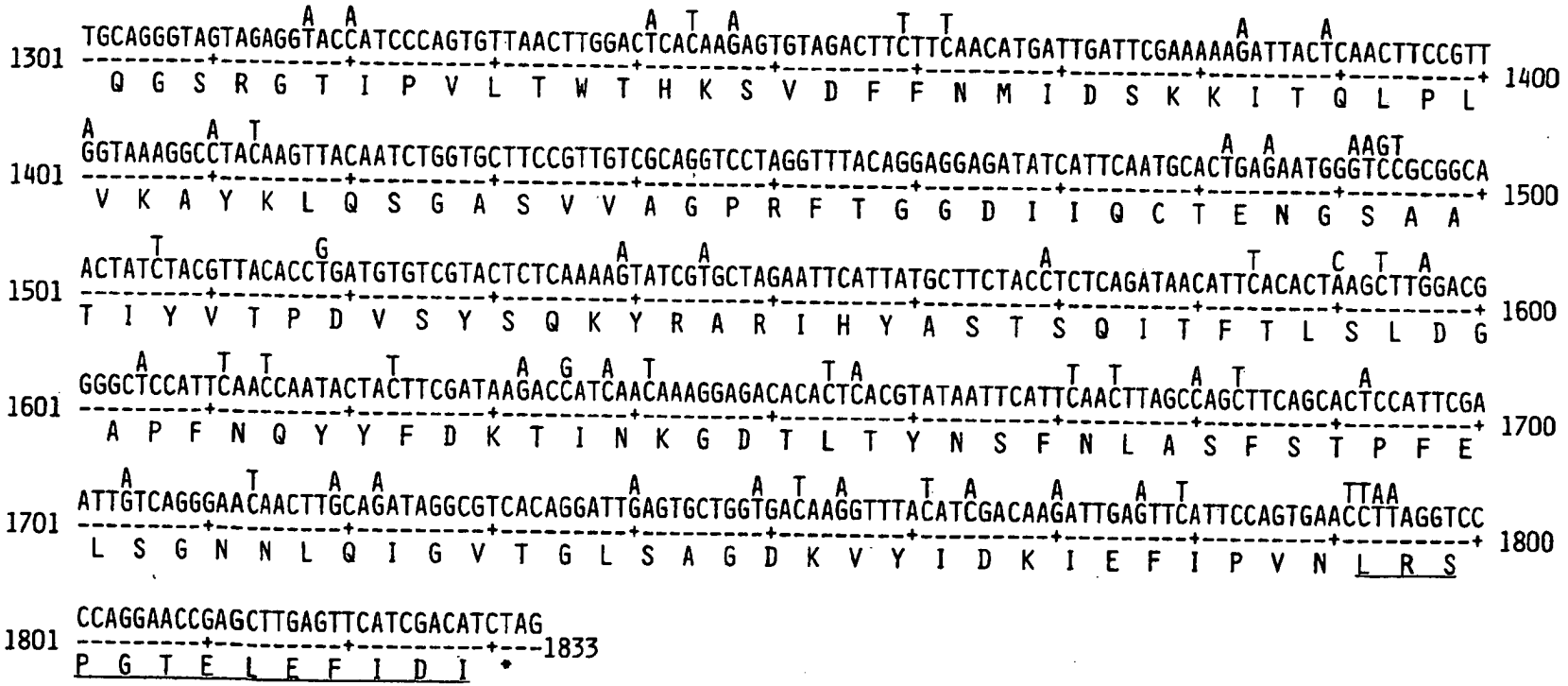

1301 TGCAGGGTAGTAGAGGTACCATCCCAGTGTTAACTTGGACTCACAAGAGTGTAGACTTCTTCAACATGATTGATTCGAAAAAGATTACTCAACTTCCGTT
Q G S R G T I P V L T W T H K S V D F F N M I D S K K I T Q L P L 1400
1401 GGTAAAGGCCTACAAGTTACAATCTGGTGCTTCCGTTGTCGCAGGTCCTAGGTTTACAGGAGGAGATATCATTCAATGCACTGAGAATGGGTCCGCGGCA
V K A Y K L Q S G A S V V A G P R F T G G D I I Q C T E N G S A A 1500
1501 ACTATCTACGTTACACCTGATGTGTCGTACTCTCAAAAGTATCGTGCTAGAATTCATTATGCTTCTACCTCTCAGATAACATTCACACTAAGCTTGGACG
T I Y V T P D V S Y S Q K Y R A R I H Y A S T S Q I T F T L S L D G 1600
1601 GGGCTCCATTCAACCAATACTACTTCGATAAGACCATCAACAAAGGAGACACACTCACGTATAATTCATTCAACTTAGCCAGCTTCAGCACTCCATTCGA
A P F N Q Y Y F D K T I N K G D T L T Y N S F N L A S F S T P F E 1700
1701 ATTGTCAGGGAACAACTTGCAGATAGGCGTCACAGGATTGAGTGCTGGTGACAAGGTTTACATCGACAAGATTGAGTTCATTCCAGTGAACCTTAGGTCC
L S G N N L Q I G V T G L S A G D K V Y I D K I E F I P V N L R S 1800
1801 CCAGGAACCGAGCTTGAGTTCATCGACATCTAG
P G T E L E F I D I * 1833

FIG. IC

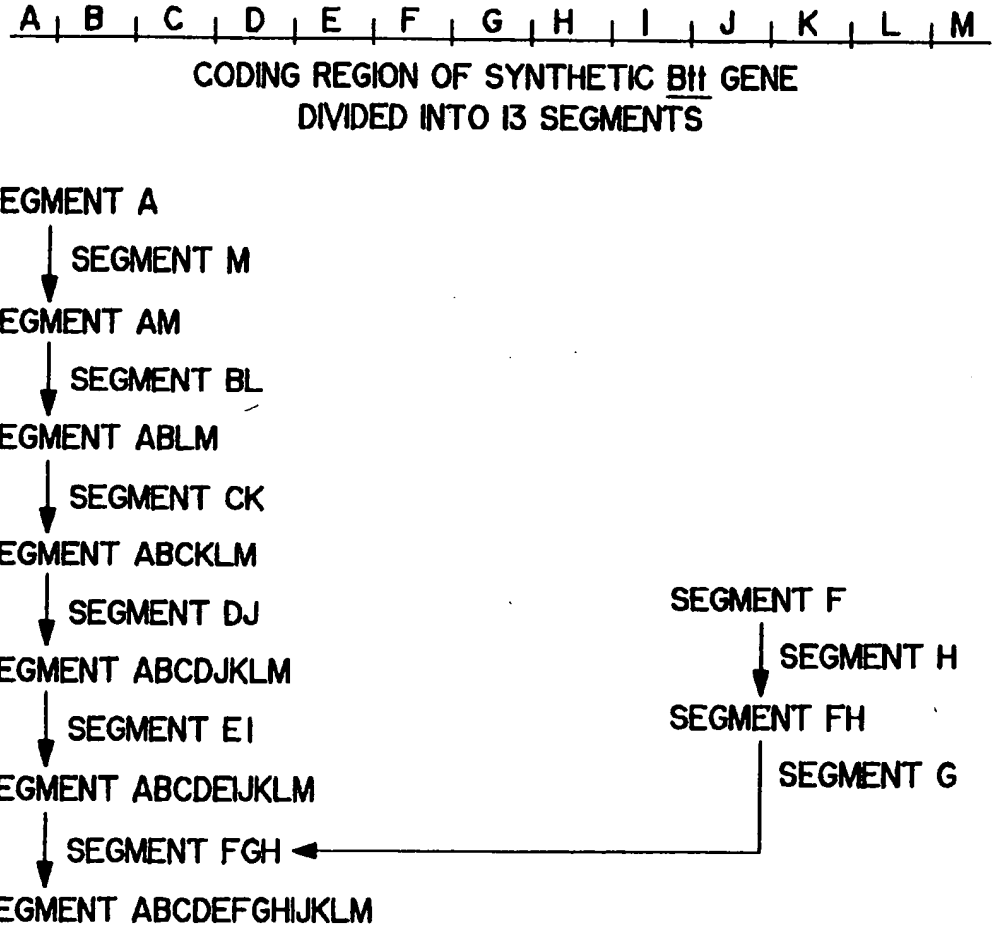

A B C D E F G H I J K L M
CODING REGION OF SYNTHETIC Bt1 GENE
DIVIDED INTO 13 SEGMENTS
SEGMENT A
SEGMENT M
SEGMENT AM
SEGMENT BL
SEGMENT ABLM
SEGMENT CK
SEGMENT ABCKLM
SEGMENT DJ
SEGMENT ABCDJKLM
SEGMENT EI
SEGMENT ABCDEIJKLM
SEGMENT FGH
SEGMENT ABCDEFGHIJKLM
SEGMENT F
SEGMENT H
SEGMENT FH
SEGMENT G

FIG. 2

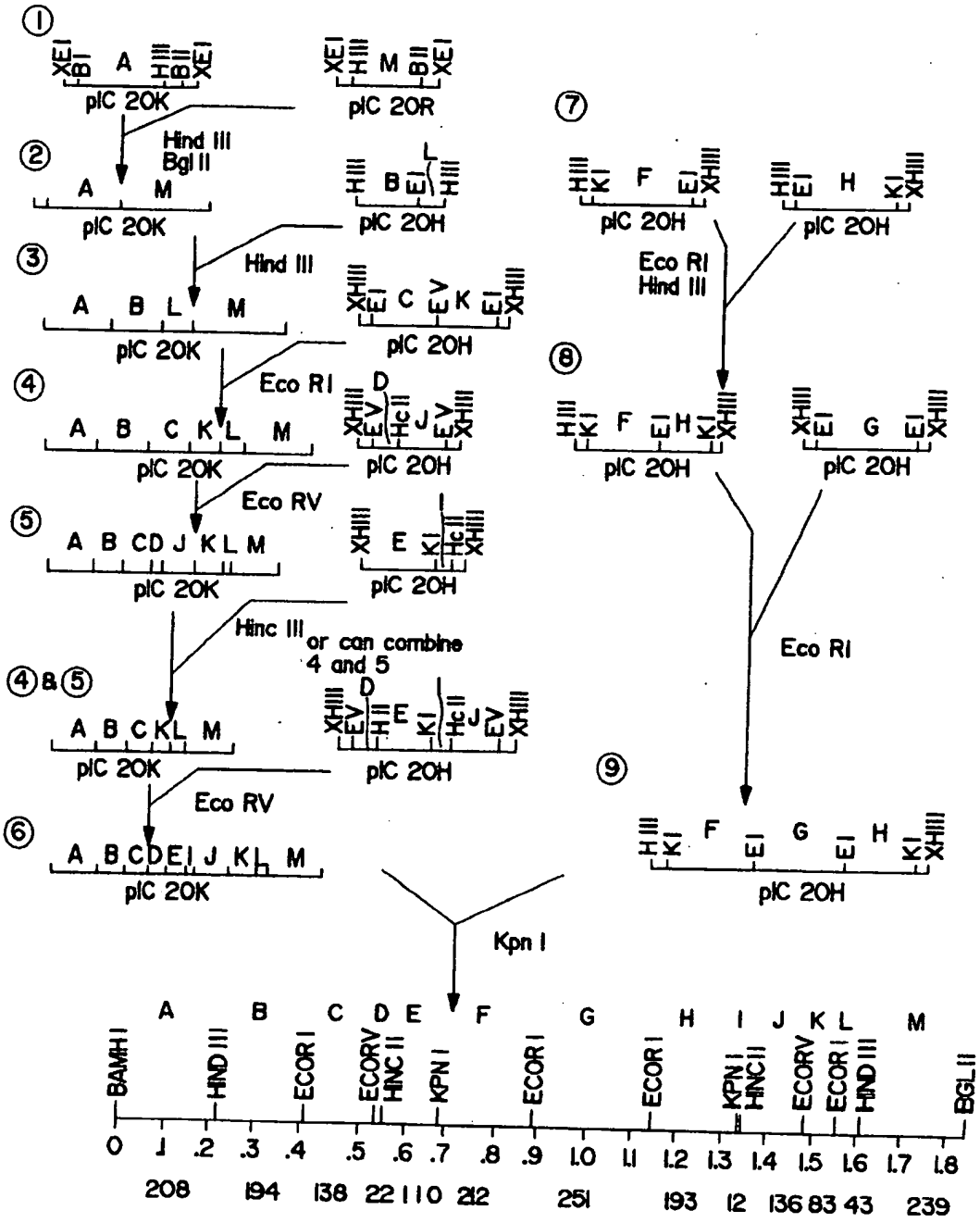

pIC 20K
pIC 20R
pIC 20H
Hind III
Bgl II
Eco RI
Eco RV
Hinc III
or can combine 4 and 5
Kpn I
BAMH I
HND III
ECOR I
ECORV
HNC II
KPN I
BGL II
208 194 138 22 110 212 251 193 12 136 83 43 239

FIG. 3

SYNTHETIC INSECTICIDAL CRYSTAL PROTEIN GENE HAVING A MODIFIED FREQUENCY OF CODON USAGE

CROSS REFERENCES TO RELATED APPLICATIONS

This is a division of application Ser. No. 08/369,835, filed Jan. 6, 1995, issued as U.S. Pat. No. 5,567,600, which is a continuation-in-part of application Ser. No. 08/057,191, filed May 3, 1993 issued as U.S. Pat. No. 5,380,831, which is a continuation of application Ser. No. 07/827,844, filed Jan. 28, 1992, now abandoned; which is a continuation of application Ser. No. 07/242,482, filed Sep. 9, 1988, now abandoned, all of which are incorporated herein by reference.

FIELD OF THE INVENTION

This invention relates to the field of bacterial molecular biology and, in particular, to genetic engineering by recombinant technology for the purpose of protecting plants from insect pests. Disclosed herein ar ethe chemical synthesis of a modified crystal protein gene from *Bacillus thuringiensis* var. *tenebrionis* (*Btt*), and the selective expression of this synthetic insecticidal gene. Also disclosed is the transfer of the cloned synthetic gene into a host microoorganism, rendering the organism capable of producing, at improved levels of expression, a protein having toxicity to insects. This invention facilitates the genetic engineering of bacteria and plants to attain desired expression levels of novel toxins having agronomic value.

BACKGROUND OF THE INVENTION

B. thuringiensis (*Bt*) is unique in its ability to produce, during the process of sporulation, proteinaceous, crystalline inclusions which are found to be highly toxic to several insect pests of agricultural importance. The crystal proteins of different *Bt* strains have a rather narrow host range and hence are used commercially as very selective biological insecticides. Numerous strains of *Bt* are toxic to lepidopteran and dipteran insects. Recently two subspecies (or varieties) of *Bt* have been reported to be pathogenic to coleopteran insects: var. *tenebrionis* (Krieg et al. (1983) Z. Angew. Entomol. 96:500–508) and var. *san diego* (Herrnstadt et al. (1986) Biotechnol. 4:305–308). Both strains produce flat, rectangular crystal inclusions and have a major crystal component of 64–68 kDa (Herrnstadt et al. supra; Bernhard (1986) FEMS Microbiol. Lett. 33:261–265).

Toxin genes from several subspecies of *Bt* have been cloned and the recombinant clones were found to be toxic to lepidopteran and dipteran insect larvae. The two coleopteran-active toxin genes have also been isolated and expressed. Herrnstadt et al. supra cloned a 5.8 kb BamHI fragment of *Bt* var. *san diego* DNA. The protein expressed in *E. coli* was toxic to *P. luteola* (Elm leaf beetle) and had a molecular weight of approximately 83 kDa. This 83 kDa toxin product from the var. *san diego* gene was larger than the 64 kDa crystal protein isolated from *Bt* var. *san diego* cells, suggesting that the *Bt* var. *san diego* crystal protein may be synthesized as a larger precursor molecule that is processed by *Bt* var. *san diego* but not by *E. coli* prior to being formed into a crystal.

Sekar et al. (1987) Proc. Nat. Acad. Sci. USA 84:7036–7040; U.S. patent application Ser. No. 108,285, filed Oct. 13, 1987 isolated the crystal protein gene from *Btt*

and determined the nucleotide sequence. This crystal protein gene was contained on a 5.9 kb BamHI fragment (pNSBF544). A subclone containing the 3 kb HindIII fragment from pNSBF544 was constructed. This HindIII fragment contains an open reading frame (ORF) that encodes a 644-amino acid polypeptide of approximately 73 kDa. Extracts of both subclones exhibited toxicity to larvae of Colorado potato beetle (*Leptinotarsa decemlineata*, a coleopteran insect). 73- and 65-kDa peptides that cross-reacted with an antiserum against the crystal protein of var. *tenebrionis* were produced on expression in *E. coli*. Sporulating var. *tenebrionis* cells contain an immunoreactive 73-kDa peptide that corresponds to the expected product from the ORF of pNSBP544. However, isolated crystals primarily contain a 65-kDa component. When the crystal protein gene was shortened at the N-terminal region, the dominant protein product obtained was the 65-kDa peptide. A deletion derivative, p544Pst-Met5, was enzymatically derived from the 5.9 kb BamHI fragment upon removal of forty-six amino acid residues from the N-terminus. Expression of the N-terminal deletion derivative, p544Pst-Met5, resulted in the production of, almost exclusively, the 65 kDa protein. Recently, McPherson et al. (1988) Biotechnology 6:61–66 demonstrated that the *Btt* gene contains two functional translational initiation codons in the same reading frame leading to the production of both the full-length protein and an N-terminal truncated form.

Chimeric toxin genes from several strains of *Bt* have been expressed in plants. Four modified *Bt2* genes from var. *berliner* 1715, under the control of the 2' promoter of the Agrobacterium TR-DNA, were transferred into tobacco plants (Vaeck et al. (1987) Nature 328:33–37). Insecticidal levels of toxin were produced when truncated genes were expressed in transgenic plants. However, the steady state mRNA levels in the transgenic plants were so low that they could not be reliably detected in Northern blot analysis and hence were quantified using ribonuclease protection experiments. *Bt* mRNA levels in plants producing the highest level of protein corresponded to ~0.0001% of the poly(A)$^+$ mRNA.

In the report by Vaeck et al. (1987) supra, expression of chimeric genes containing the entire coding sequence of *Bt2* were compared to those containing truncated *Bt2* genes. Additionally, some T-DNA constructs included a chimeric NPTII gene as a marker selectable in plants, whereas other constructs carried translational fusions between fragments of *Bt2* and the NPTII gene. Insecticidal levels of toxin were produced when truncated *Bt2* genes or fusion constructs were expressed in transgenic plants. Greenhouse grown plants produced ~0.02% of the total soluble protein as the toxin, or 3 μg of toxin per g. fresh leaf tissue and, even at five-fold lower levels, showed 100% mortality in six-day feeding assays. However, no significant insecticidal activity could be obtained using the intact *Bt2* coding sequence, despite the fact that the-same promoter was used to direct its expression. Intact *Bt2* protein and RNA yields in the transgenic plant leaves were 10–50 times lower than those for the truncated *Bt2* polypeptides or fusion proteins.

Barton et al. (1987) Plant Physiol. 85:1103–1109 showed expression of a *Bt* protein in a system containing a 35S promoter, a viral (TMV) leader sequence, the *Bt* HD-1 4.5 kb gene (encoding a 645 amino acid protein followed by two proline residues) and a nopaline synthase (nos) poly(A)+ sequence. Under these conditions expression was observed for *Bt* mRNA at levels up to 47 pg/20 μg RNA and 12 ng/mg plant protein. This amount of *Bt* protein in plant tissue produced 100% mortality in two days. This level of expres-

sion still represents a low level of mRNA (2.5×10^{-4}%) and protein (1.2×10^{-3}%).

Various hybrid proteins consisting of N-terminal fragments of increasing length of the *Bt*2 protein fused to NPTII were produced in *E. coli* by Hofte et al. (1988) FEBS Lett. 226:364–370. Fusion proteins containing the first 607 amino acids of *Bt*2 exhibited insect toxicity; fusion proteins not containing this minimum N-terminal fragment were non-toxic. Appearance of NPTII activity was not dependent upon the presence of insecticidal activity; however, the conformation of the *Bt*2 polypeptide appeared to exert an important influence on the enzymatic activity of the fused NPTII protein. This study did suggest that the global 3-D structure of the *Bt*2 polypeptide is disturbed in truncated polypeptides.

A number of researchers have attempted to express plant genes in yeast (Neill et al. (1987) Gene 55:303–317; Rothstein et al. (1987) Gene 55:353–356; Coraggio et al. (1986) EMBO J. 5:459–465) and *E. coli* (Fuzakawa et al. (1987) FEBS Lett. 224:125–127; Vies et al. (1986) EMBO J. 5:2439–2444; Gatenby et al. (1987) Eur. J. Biochem. 168:227–231). In the case of wheat α-gliadin (Neill et al. (1987) supra), α-amylase (Rothstein et al. (1987) supra) genes, and maize zein genes (Coraggio et al. (1986) supra) in yeast, low levels of expression have been reported. Neill et al. have suggested that the low levels of expression of α-gliadin in yeast may be due in part to codon usage bias, since α-gliadin codons for Phe, Leu, Ser, Gly, Tyr and especially Glu do not correlate well with the abundant yeast isoacceptor tRNAs. In *E. coli* however, soybean glycinin A2 (Fuzakawa et al. (1987) supra) and wheat RuBPC SSU (Vies et al. (1986) supra; Gatenby et al. (1987) supra) are expressed adequately.

Not much is known about the makeup of tRNA populations in plants. Viotti et al. (1978) Biochim. Biophys. Acta 517:125–132 report that maize endosperm actively synthesizing zein, a storage protein rich in glutamine, leucine, and alanine, is characterized by higher levels of accepting activity for these three amino acids than are maize embryo tRNAs. This may indicate that the tRNA population of specific plant tissues may be adapted for optimum translation of highly expressed proteins such as zein. To our knowledge, no one has experimentally altered codon bias in highly expressed plant genes to determine possible effects of the protein translation in plants to check the effects on the level of expression.

SUMMARY OF THE INVENTION

It is the overall object of the present invention to provide a means for plant protection against insect damage. The invention disclosed herein comprises a chemically synthesized gene encoding an insecticidal protein which is functionally equivalent to a native insecticidal protein of *Bt*. This synthetic gene is designed to be expressed in plants at a level higher than a native *Bt* gene. It is preferred that the synthetic gene be designed to be highly expressed in plants as defined herein. Preferably, the synthetic gene is at least approximately 85% homologous to an insecticidal protein gene of *Bt*.

It is a particular object of this invention to provide a synthetic structural gene coding for an insecticidal protein from *Btt* having, for example, the nucleotide sequences presented in FIG. 1 and spanning nucleotides 1 through 1793 or spanning nucleotide 1 through 1833 with functional equivalence.

In designing synthetic *Btt* genes of this invention for enhanced expression in plants, the DNA sequence of the native *Btt* structural gene is modified in order to contain codons preferred by highly expressed plant genes, to attain an A+T content in nucleotide base composition substantially that found in plants, and also preferably to form a plant initiation sequence, and to eliminate sequences that cause destabilization, inappropriate polyadenylation, degradation and termination of RNA and to avoid sequences that constitute secondary structure hairpins and RNA splice sites. In the synthetic genes, codons used to specify a given amino acid are selected with regard to the distribution frequency of codon usage employed in highly expressed plant genes to specify that amino acid. As is appreciated by those skilled in the art, the distribution frequency of codon usage utilized in the synthetic gene is a determinant of the level of expression. Hence, the synthetic gene is designed such that its distribution frequency of codon usage deviates, preferably, no more than 25% from that of highly expressed plant genes and, more preferably, no more than about 10%. In addition, consideration is given to the percentage G+C content of the degenerate third base (monocotyledons appear to favor G+C in this position, whereas dicotyledons do not). It is also recognized that the XCG nucleotide is the least preferred codon in dicots whereas the XTA codon is avoided in both monocots and dicots. The synthetic genes of this invention also preferably have CG and TA doublet avoidance indices as defined in the Detailed Description closely approximating those of the chosen host plant. More preferably these indices deviate from that of the host by no more than about 10–15%.

Assembly of the *Bt* gene of this invention is performed using standard technology known to the art. The *Btt* structural gene designed for enhanced expression in plants of the specific embodiment is enzymatically assembled within a DNA vector from chemically synthesized oligonucleotide duplex segments. The synthetic *Bt* gene is then introduced into a plant host cell and expressed by means known to the art. The insecticidal protein produced upon expression of the synthetic *Bt* gene in plants is functionally equivalent to a native *Bt* crystal protein in having toxicity to the same insects.

BRIEF DESCRIPTION OF THE FIGURES

FIGS. 1A and 1B present the nucleotide sequence for the synthetic *Btt* gene. Where different, the native sequence as found in p544Pst-Met5 is shown above. Changes in amino acids (underlined) occur in the synthetic sequence with alanine replacing threonine at residue 2 and leucine replacing the stop at residue 596 followed by the addition of 13-amino acids at the C-terminus.

FIG. 2 represents a simplified scheme used in the construction of the synthetic *Btt* gene. Segments A through M represent oligonucleotide pieces annealed and ligated together to form DNA duplexes having unique splice sites to allow specific enzymatic assembly of the DNA segments to give the desired gene.

FIG. 3 is a schematic diagram showing the assembly of oligonucleotide segments in the construction of a synthetic *Btt* gene. Each segment (A through M) is built from oligonucleotides of different sizes, annealed and liated to form the desired DNA segment.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO. 1 is the native DNA gene sequence corresponding to FIG. 1.

SEQ ID NO. 2 is the corresponding amino acid sequence encoded by SEQ ID NO. 1.

SEQ ID NO. 3 is the synthetic DNA gene sequence corresponding to FIG. 1.

SEQ ID NO. 4 is the corresponding amino acid sequence encoded by SEQ ID NO. 3.

SEQ ID NO. 5 is the sequence of a synthetic DNA linker described in Example 1(i).

SEQ ID NO. 6 is the sequence of the 5' plant consensus splice site found in Example 1(iii)(d).

SEQ ID NO. 7 is the sequence of the 3' plant consensus splice site found in Example 1(iii)(d).

SEQ ID NO. 8 is the nucleotide sequence of Segment A found in Table 4.

SEQ ID NO. 9 is the nucleotide sequence of Segment M found in Table 5.

DETAILED DESCRIPTION OF THE INVENTION

The following definitions are provided in order to provide clarity as to the intent or scope of their usage in the Specification and claims.

Expression refers to the transcription and translation of a structural gene to yield the encoded protein. The synthetic *Bt* genes of the present invention are designed to be expressed at a higher level in plants than the corresponding native *Bt* genes. As will be appreciated by those skilled in the art, structural gene expression levels are affected by the regulatory DNA sequences (promoter, polyadenylation sites, enhancers, etc.) employed and by the host cell in which the structural gene is expressed. Comparisons of synthetic *Bt* gene expression and native *Bt* gene expression must be made employing analogous regulatory sequences and in the same host cell. It will also be apparent that analogous means of assessing gene expression must be employed in such comparisons.

Promoter refers to the nucleotide sequences at the 5' end of a structural gene which direct the initiation of transcription. Promoter sequences are necessary, but not always sufficient, to drive the expression of a downstream gene. In prokaryotes, the promoter drives transcription by providing binding sites to RNA polymerases and other initiation and activation factors. Usually promoters drive transcription preferentially in the downstream direction, although promotional activity can be demonstrated (at a reduced level of expression) when the gene is placed upstream of the promoter. The level of transcription is. regulated by promoter sequences. Thus, in the construction of heterologous promoter/structural gene combinations, the structural gene is placed under the regulatory control of a promoter such that the expression of the gene is controlled by promoter sequences. The promoter is positioned preferentially upstream to the structural gene and at a distance from the transcription start site that approximates the distance between the promoter and the gene it controls in its natural setting. As is known in the art, some variation in this distance can be tolerated without loss of promoter function.

A gene refers to the entire DNA portion involved in the synthesis of a protein. A gene embodies the structural or coding portion which begins at the 5' end from the translational start codon (usually ATG) and extends to the stop (TAG, TGA or TAA) codon at the 3' end. It also contains a promoter region, usually located 5' or upstream to the structural gene, which initiates and regulates the expression of a structural gene. Also included in a gene are the 3' end and poly(A)+ addition sequences.

Structural gene is that portion of a gene comprising a DNA segment encoding a protein, polypeptide or a portion thereof, and excluding the 5' sequence which drives the initiation of transcription. The structural gene nay be one which is normally found in the cell or one which is not normally found in the cellular location wherein it is introduced, in which case it is termed a heterologous gene. A heterologous gene may be derived in whole or in part from any source know to the art, including a bacterial genome or episome, eukaryotic, nuclear or plasmid DNA, cDNA, viral DNA or chemically synthesized DNA. A structural gene may contain one or more modifications in either the coding or the untranslated regions which could affect the biological activity or the chemical structure of the expression product, the rate of expression or the manner of expression control. Such modifications include, but are not limited to, mutations, insertions, deletions and substitutions of one or more nucleotides. The structural gene may constitute an uninterrupted coding sequence or it may include one or more introns, bounded by the appropriate splice junctions. The structural gene may be a composite of segments derived from a plurality of sources, naturally occurring or synthetic. The structural gene may also encode a fusion protein.

Synthetic gene refers to a DNA sequence of a structural gene that is chemically synthesized in its entirety or for the greater part of the coding region. As exemplified herein, oligonucleotide building blocks are synthesized using procedures known to those skilled in the art and are ligated and annealed to form gene segments which are then enzymatically assembled to construct the entire gene. As is recognized by those skilled in the art, functionally and structurally equivalent genes to the synthetic genes described herein may be prepared by site-specific mutagenesis or other related methods used in the art.

Transforming refers to stably introducing a DNA segment carrying a functional gene into an organism that did not previously contain that gene.

Plant tissue includes differentiated and undifferentiated tissues of plants, including but not limited to, roots, shoots, leaves, pollen, seeds, tumor tissue and various forms of cells in culture, such as single cells, protoplasts, embryos and callus tissue. The plant tissue may be in planta or in organ, tissue or cell culture.

Plant cell as used herein includes plant cells in planta and plant cells and protoplasts in culture.

Homology refers to identity or near identity of nucleotide or amino acid sequences. As is understood in the art, nucleotide mismatches can occur at the third or wobble base in the codon without causing amino acid substitutions in the final polypeptide sequence. Also, minor nucleotide modifications (e.g., substitutions, insertions or deletions) in certain regions of the gene sequence can be tolerated and considered insignificant whenever such modifications result in changes in amino acid sequence that do not alter functionality of the final product. It has been shown that chemically synthesized copies of whole, or parts of, gene sequences can replace the corresponding regions in the natural gene without loss of gene function. Homologs of specific DNA sequences may be identified by those skilled in the art using the test of cross-hybridization of nucleic acids under conditions of stringency as is well understood in the art (as described in Hames and Higgens (eds.) (1985) *Nucleic Acid Hybridization*, IRL Press, Oxford, UK). Extent of homology is often measured in terms of percentage of identity between the sequences compared.

Functionally equivalent refers to identity or near identity of function. A synthetic gene product which is toxic to at least one of the same insect species as a natural *Bt* protein is considered functionally equivalent thereto. As exemplified herein, both natural and synthetic *Btt* genes encode 65 kDa

insecticidal proteins having essentially identical amino acid sequences and having toxicity to coleopteran insects. The synthetic *Bt* genes of the present invention are not considered to be functionally equivalent to native *Bt* genes, since they are expressible at a higher level in plants than native *Bt* genes.

Frequency of preferred codon usage refers to the preference exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. To determine the frequency of usage of a particular codon in a gene, the number of occurrences of that codon in the gene is divided by the total number of occurrences of all codons specifying the same amino acid in the gene. Table 1, for example, gives the frequency of codon usage for *Bt* genes, which was obtained by analysis of four *Bt* genes whose sequences are publicly available. Similarly, the frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell. It is preferable that this analysis be limited to genes that are highly expressed by the host cell. Table 1, for example, gives the frequency of codon usage by highly expressed genes exhibited by dicotyledonous plants, and monocotyledonous plants. The dicot codon usage was calculated using 154 highly expressed coding sequences obtained from Genbank which are listed in Table 1. Monocot codon usage was calculated using 53 monocot nuclear gene coding sequences obtained from Genbank and listed in Table 1, located in Example 1.

When synthesizing a gene for improved expression in a host call it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

The percent deviation of the frequency of preferred codon usage for a synthetic gene from that employed by a host cell is calculated first by determining the percent deviation of the frequency of usage of a single codon from that of the host cell followed by obtaining the average deviation over all codons. As defined herein this calculation includes unique codons (i.e., ATG and TGG). The frequency of preferred codon usage of the synthetic *Btt* gene, whose sequence is given in FIGS. 1A–C, is given in Table 1. The frequency of preferred usage of the codon 'GTA' for valine in the synthetic gene (0.10) deviates from that preferred by dicots (0.12) by 0.02/0.12=0.167 or 16.7%. The average deviation over all amino acid codons of the *Btt* synthetic gene codon usage from that of dicot plants is 7.8%. In general terms the overall average deviation of the codon usage of a synthetic gene from that of a host cell is calculated using the equation

$$A = \sum_{n=1}^{Z} \frac{\frac{X_n - Y_n}{X_n} \times 100}{Z}$$

where X_n=frequency of usage for codon n in the host cell; Y_n=frequency of usage for codon n in the synthetic gene. Where n represents an individual codon that specifies an amino acid, the total number of codons is Z, which in the preferred embodiment is 61. The overall deviation of the frequency of codon usage, A, for all amino acids should preferably be less than about 25%, and more preferably less than about 10%.

Derived from is used to mean taken, obtained, received, traced, replicated or descended from a source (chemical and/or biological). A derivative may be produced by chemical or biological manipulation (including but not limited to substitution, addition, insertion, deletion, extraction, isolation, mutation and replication) of the original source.

Chemically synthesized, as related to a sequence of DNA, means that the component nucleotides were assembled in vitro. Manual chemical synthesis of DNA may be accomplished using well established procedures (Caruthers, M. (1983) in *Methodology of DNA and RNA Sequencing*, Weissman (ed.), Praeger Publishers, New York, Chapter 1), or automated chemical synthesis can be performed using one of a number of commercially available machines.

The term, designed to be highly expressed as used herein refers to a level of expression of a designed gene wherein the amount of its specific mRNA transcripts produced is sufficient to be quantified in Northern blots and, thus, represents a level of specific mRNA expressed corresponding to greater than or equal to approximately 0.001% of the poly(A)+ mRNA. To date, natural *Bt* genes are transcribed at a level wherein the amount of specific mRNA produced is insufficient to be estimated using the Northern blot technique. However, in the present invention, transcription of a synthetic *Bt* gene designed to be highly expressed not only allows quantification of the specific mRNA transcripts produced but also results in enhanced expression of the translation product which is measured in insecticidal bioassays.

Crystal protein or insecticidal crystal protein or crystal toxin refers to the major protein component of the parasporal crystals formed in strains of *Bt*. This protein component exhibits selective pathogenicity to different species of insects. The molecular size of the major protein isolated from parasporal crystals varies depending on the strain of *Bt* from which it is derived. Crystal proteins having molecular weights of approximately 132, 65, and 28 kDa have been reported. It has been shown that the approximately 132 kDa protein is a protoxin that is cleaved to form an approximately 65 kDa toxin.

The crystal protein gene refers to the DNA sequence encoding the insecticidal crystal protein in either full length protoxin or toxin form, depending on the strain of *Bt* from which the gene is derived.

The authors of this invention observed that expression in plants of *Bt* crystal protein mRNA occurs at levels that are not routinely detectable in Northern blots and that low levels of *Bt* crystal protein expression correspond to this low level of mRNA expression. It is preferred for exploitation of these genes as potential biocontrol methods that the level of expression of *Bt* genes in plant cells be improved and that the stability of *Bt* mRNA in plants be optimized. This will allow greater levels of *Bt* mRNA to accumulate and will result in an increase in the amount of insecticidal protein in plant tissues. This is essential for the control of insects that are relatively resistant to *Bt* protein.

Thus, this invention is based on the recognition that expression levels of desired, recombinant insecticidal protein in transgenic plants can be improved via increased expression of stabilized mRNA transcripts; and that, conversely, detection of these stabilized RNA transcripts may be utilized to measure expression of translational product (protein). This invention provides a means of resolving the problem of low expression of insecticidal protein RNA in plants and, therefore, of low protein expression through the use of an improved, synthetic gene specifying an insecticidal crystal protein from *Bt*.

Attempts to improve the levels of expression of *Bt* genes in plants have centered on comparative studies evaluating parameters such as gene type, gene length, choice of promoters, addition of plant viral untranslated RNA leader, addition of intron sequence and modification of nucleotides

surrounding the initiation ATG codon. To date, changes in these parameters have not led to significant enhancement of *Bt* protein expression in plants. Applicants find that, surprisingly, to express *Bt* proteins at the desired level in plants, modifications in the coding region of the gene were effective. Structural-function relationships can be studied using site-specific mutagenesis by replacement of restriction fragments with synthetic DNA duplexes containing the desired nucleotide changes (Lo et al. (1984) Proc. Natl. Acad. Sci. 81:2285–2289). However, recent advances in recombinant DNA technology now make it feasible to chemically synthesize an entire gene designed specifically for a desired function. Thus, the *Btt* coding region was chemically synthesized, modified in such a way as to improve its expression in plants. Also, gene synthesis provides the opportunity to design the gene so as to facilitate its subsequent mutagenesis by incorporating a number of appropriately positioned restriction endonuclease sites into the gene.

The present invention provides a synthetic *Bt* gene for a crystal protein toxic to an insect. As exemplified herein, this protein is toxic to coleopteran insects. To the end of improving expression of this insecticidal protein in plants, this invention provides a DNA segment homologous to a *Btt* structural gene and, as exemplified herein, having approximately 85% homology to the *Btt* structural gene in p544Pst-Met5. In this embodiment the structural gene encoding a *Btt* insecticidal protein is obtained through chemical synthesis of the coding region. A chemically synthesized gene is used in this embodiment because it best allows for easy and efficacious accommodation of modifications in nucleotide sequences required to achieve improved levels of cross-expression.

Today, in general, chemical synthesis is a preferred method to obtain a desired modified gene. However, to date, no plant protein gene has been chemically synthesized nor has any synthetic gene for a bacterial protein been expressed in plants. In this invention, the approach adopted for synthesizing the gene consists of designing an improved nucleotide sequence for the coding region and assembling the gene from chemically synthesized oligonucleotide segments. In designing the gene, the coding region of the naturally-occurring gene, preferably from the *Btt* subclone, p544Pst-Met5, encoding a 65 kDa polypeptide having coleoperan toxicity, is scanned for possible modifications which would result in improved expression of the synthetic gene in plants. For example, to optimize the efficiency of translation, codons preferred in highly expressed proteins of the host cell are utilized.

Bias in codon choice within genes in a single species appears related to the level of expression of the protein encoded by that gene. Codon bias is most extreme in highly expressed proteins of *E. coli* and yeast. In these organisms, a strong positive correlation has been reported between the abundance of an isoaccepting tRNA species and the favored synonymous codon. In one group of highly expressed proteins in yeast, over 96% of the amino acids are encoded by only 25 of the 61 available codons (Bennetzen and Hall (1982) J. Biol. Chem. 257:3026–3031). These 25 codons are preferred in all sequenced yeast genes, but the degree of preference varies with the level of expression of the genes. Recently, Hoekema and colleagues (1987) Mol. Cell. Biol. 7:2914–2924 reported that replacement of these 25 preferred codons by minor codons in the 5' end of the highly expressed yeast gene PGK1 results in a decreased level of both protein and mRNA. They concluded that biased codon choice in highly expressed genes enhances translation and is required for maintaining mRNA stability in yeast. Without doubt, the degree of codon bias is an important factor to consider when engineering high expression of heterologous genes in yeast and other systems.

Experimental evidence obtained from point mutations and deletion analysis has indicated that in eukaryotic genes specific sequences are associated with post-transcriptional processing, RNA destabilization, translational termination, intron splicing and the like. These are preferably employed in the synthetic genes of this invention. In designing a bacterial gene for expression in plants, sequences which interfere with the efficacy of gene expression are eliminated.

In designing a synthetic gene, modifications in nucleotide sequence of the coding region are made to modify the A+T content in DNA base composition of the synthetic gene to reflect that normally found in genes for highly expressed proteins native to the host cell. Preferably the A+T content of the synthetic gene is substantially equal to that of said genes for highly expressed proteins. In genes encoding highly expressed plant proteins, the A+T content is approximately 55%. It is preferred that the synthetic gene have an A+T content near this value, and not sufficiently high as to cause destabilization of RNA and, therefore, lower the protein expression levels. More preferably, the A+T content is no more than about 60% and most preferably is about 55%. Also, for ultimate expression in plants, the synthetic gene nucleotide sequence is preferably modified to form a plant initiation sequence at the 5' end of the coding region. In addition, particular attention is preferably given to assure that unique restriction sites are placed in strategic positions to allow efficient assembly of oligonucleotide segments during construction of the synthetic gene and to facilitate subsequent nucleotide modification. As a result of these modifications in coding region of the native *Bt* gene, the preferred synthetic gene is expressed in plants at an enhanced level when compared to that observed with natural *Bt* structural genes.

In specific embodiments, the synthetic *Bt* gene of this invention encodes a *Btt* protein toxic to coleopteran insects. Preferably, the toxic polypeptide is about 598 amino acids in length, is at least 75% homologous to a *Btt* polypeptide, and, as exemplified herein, is essentially identical to the protein encoded by p544Pst-Met5, except for replacement of threonine by alanine at residue 2. This amino acid substitution results as a consequence of the necessity to introduce a guanine base at position +4 in the coding sequence.

In designing the synthetic gene of this invention, the coding region from the *Btt* subclone, p544Pst-Met5, encoding a 65 kDa polypeptide having coleopteran toxicity, is scanned for possible modifications which would result in improved expression of the synthetic gene in plants. For example, in preferred embodiments, the synthetic insecticidal protein is strongly expressed in dicot plants, e.g., tobacco, tomato, cotton, etc., and hence, a synthetic gene under these conditions is designed to incorporate to advantage codons used preferentially by highly expressed dicot proteins. In embodiments where enhanced expression of insecticidal protein is desired in a monocot, codons preferred by highly expressed monocot proteins (given in Table 1) are employed in designing the synthetic gene.

In general, genes within a taxonomic group exhibit similarities in codon choice, regardless of the function of these genes. Thus an estimate of the overall use of the genetic code by a taxonomic group can be obtained by summing codon frequencies of all its sequenced genes. This species-specific codon choice is reported in this invention from analysis of 208 plant genes. Both monocot and dicot plants are analyzed

individually to determine whether these broader taxonomic groups are characterized by different patterns of synonymous codon preference. The 208 plant genes included in the codon analysis code for proteins having a wide range of functions and they represent 6 monocot and 36 dicot species. These proteins are present in different plant tissues at varying levels of expression.

In this invention it is shown that the relative use of synonymous codons differs between the monocots and the dicots. In general, the most important factor in discriminating between monocot and dicot patterns of codon usage is the percentage G+C content of the degenerate third base. In monocots, 16 of 18 amino acids favor G+C in this position, while dicots only favor G+C in 7 of 18 amino acids.

The G ending codons for Thr, Pro, Ala and Ser are avoided in both inonocots and dicots because they contain C in codon position II. The CG dinucleotide is strongly avoided in plants (Boudraa (1987) Genet. Sel. Evol. 19:143–154) and other eukaryotes (Grantham et al. (1985) Bull. Inst. Pasteur 83:95–148), possibly due to regulation involving methylation. In dicots, XCG is always the least favored codon, while in monocots this is not the case. The doublet TA is also avoided in codon positions II and III in most eukaryotes, and this is true of both nonocots and dicots.

Grantham and colleagues (1986) Oxford Surveys in Evol. Biol. 3:48–81 have developed two codon choice indices to quantify CG and TA doublet avoidance in codon positions II and III. XCG/XCC is the ratio of codons having C as base II of G-ending to C-ending triplets, while XTA/XTT is the ratio of A-ending to T-ending triplets with T as the second base. These indices have been calculated for the plant data in this paper (Table 2) and support the conclusion that monocot and dicot species differ in their use of these dinucleotides.

TABLE 2

Avoidance of CG and TA doublets in codons position II–III. XCG/XCC and XTA/XAA values are multiplied by 100.

Group	Plants	Di-cots	Mono-cots	Maize	Soy-bean	RuBPC SSU	CAB
XCG/XCC	40	30	61	67	37	18	22
XTA/XTT	37	35	47	43	41	9	13

RuBPC SSU = ribulose 1,5 bisphosphate small subunit
CAB = chlorophyll a/b binding protein

Additionally, for two species, soybean and maize, species-specific codon usage profiles were calculated (not shown). The maize codon usage pattern resembles that of monocots in general, since these sequences represent over half of the monocot sequences available. The codon profile of the maize subsample is even more strikingly biased in its preference for G+C in codon position III. On the other hand, the soybean codon usage pattern is almost identical to the general dicot pattern, even though it represents a much smaller portion of the entire dicot sample.

In order to determine whether the coding strategy of highly expressed genes such as the ribulose 1,5 bisphosphate small subunit (RuBPC SSU) and chlorophyll a/b binding protein (CAB) is more biased than that of plant genes in general, codon usage profiles for subsets of these genes (19 and 17 sequences, respectively) were calculated (not shown). The RuBPC SSU and CAB pooled samples are characterized by stronger avoidance of the codons XCG and XTA than in the larger monocot and dicot samples (Table 2). Although most of the genes in these subsamples are dicot in origin (17/19 and 15/17), their codon profile resembles that of the monocots in that G+C is utilized in the degenerate base III.

The use of pooled data for highly expressed genes may obscure identification of species-specific patterns in codon choice. Therefore, the codon choices of individual genes for RuBPC SSU and CAB were tabulated. The preferred codons of the maize and wheat genes for RuBPC SSU and CAB are more restricted in general than are those of the dicot species. This is in agreement with Matsuoka et al. (1987) J. Biochem. 102:673–676) who noted the extreme codon bias of the maize RuBPC SSU gene as well as two other highly expressed genes in maize leaves, CAB and phosphoenolpyruvate carboxylase. These genes almost completely avoid the use of A+T in codon position III, although this codon bias was not as pronounced in non-leaf proteins such as alcohol dehydrogenase, zein 22 kDa sub-unit, sucrose synthetase and ATP/ADP translocator. Since the wheat SSU and CAB genes have a similar pattern of codon preference, this may reflect a common monocot pattern for these highly expressed genes in leaves. The CAB gene for Lemna and the RUBPC SSU genes for Chlamdomonas share a similar extreme preference for G+C in codon position III. In dicot CAB genes, however, A+T degenerate bases are preferred by some synonymous codons (e.g., GCT for Ala, CTT for Leu, GGA and GGT for Gly). In general, the G+C preference is less pronounced for both RuBPC SSU and CAB genes in dicots than in monocots.

In designing a synthetic gene for expression in plants, attempts are also made to eliminate sequences which interfere with the efficacy of gene expression. Sequences such as the plant polyadenylation signals, e.g., AATAAA, polymerase II termination sequence, e.g., $CAN_{(7\text{-}9)}AGTNNAA$, UCUUCGG hairpins and plant consensus splice sites are highlighted and, if present in the native *Btt* coding sequence, are modified so as to eliminate potentially deleterious sequences.

Modifications in nucleotide sequence of the *Btt* coding region are also preferably made to reduce the A+T content in DNA base composition. The *Btt* coding region has an A+T content of 64%, which is about 10% higher than that found in a typical plant coding region. Since A+T-rich regions typify plant intergenic regions and plant regulatory regions, it is deemed prudent to reduce the A+T content. The synthetic *Btt* gene is designed to have an A+T content of 55%, in keeping with values usually found in plants.

Also, a single modification (to introduce guanine in lieu of adenine) at the fourth nucleotide position in the *Btt* coding sequence is made in the preferred embodiment to form a sequence consonant with that believed to function as a plant initiation sequence (Taylor et al. (1987) Mol. Gen. Genet. 210:572–577) in optimization of expression. In addition, in exemplifying this invention thirty-nine nucleotides (thirteen codons) are added to the coding region of the synthetic gene in an attempt to stabilize primary transcripts. However, it appears that equally stable transcripts are obtained in the absence of this extension polypeptide containing thirty-nine nucleotides.

Not all of the above-mentioned modifications of the natural *Bt* gene must be made in constructing a synthetic *Bt* gene in order to obtain enhanced expression. For example, a synthetic gene may be synthesized for other purposes in addition to that of achieving enhanced levels of expression. Under these conditions, the original sequence of the natural *Bt* gene may be preserved within a region of DNA corresponding to one or more, but not all, segments used to construct the synthetic gene. Depending on the desired purpose of the gene, modification may encompass substitution of one or more, but not all, of the oligonucleotide segments used to construct the synthetic gene by a corresponding region of natural *Bt* sequence.

As is known to those skilled in the art of synthesizing genes (Mandecki et al. (1985) Proc. Natl. Acad. Sci. 82:3543–3547; Feretti et al. (1986) Proc. Natl. Acad. Sci. 83:599–603), the DNA sequence to be synthesized is divided into segment lengths which can be synthesized conveniently and without undue complication. As exemplified herein, in preparing to synthesize the *Btt* gene, the coding region is divided into thirteen segments (A–M). Each segment has unique restriction sequences at the cohesive ends. Segment A, for example, is 228 base pairs in length and is constructed from six oligonucleotide sections, each containing approximately 75 bases. Single-stranded oligonucleotides are annealed and ligated to form DNA segments. The length of the protruding cohesive ends in complementary oligonucleotide segments is four to five residues. In the strategy evolved for gene synthesis, the sites designed for the joining of oligonucleotide pieces and DNA segments are different from the restriction sites created in the gene.

In the specific embodiment, each DNA segment is cloned into a pIC-20 vector for amplification of the DNA. The nucleotide sequence of each fragment is determined at this stage by the dideoxy method using the recombinant phage DNA as templates and selected synthetic oligonucleotides as primers.

As exemplified herein and illustrated schematically in FIGS. **3** and **4**, each segment individually (e.g., segment M) is excised at the flanking restriction sites from its cloning vector and spliced into the vector containing segment A. Most often, segments are added as a paired segment instead of as a single segment to increase efficiency. Thus, the entire gene is constructed in the original plasmid harboring segment A. The nucleotide sequence of the entire gene is determined and found to correspond exactly to that shown in FIGS. 1A–C.

In preferred embodiments the synthetic *Btt* gene is expressed in plants at an enhanced level when compared to that observed with natural *Btt* structural genes. To that end, the synthetic structural gene is combined with a promoter functional in plants, the structural gene and the promoter region being in such position and orientation with respect to each other that the structural gene can be expressed in a cell in which the promoter region is active, thereby forming a functional gene. The promoter regions include, but are not limited to, bacterial and plant promoter regions. To express the promoter region/structural gene combination, the DNA segment carrying the combination is contained by a cell. Combinations which include plant promoter regions are contained by plant cells, which, in turn, may be contained by plants or seeds. Combinations which include bacterial promoter regions are contained by bacteria, e.g., *Bt* or *E. coli*. Those in the art will recognize that expression in types of micro-organisms other than bacteria may in some circumstances be desirable and, given the present disclosure, feasible without undue experimentation.

The recombinant DNA molecule carrying a synthetic structural gene under promoter control can be introduced into plant tissue by any means known to those skilled in the art. The technique used for a given plant species or specific type of plant tissue depends on the known successful techniques. As novel means are developed for the stable insertion of foreign genes into plant cells and for manipulating the modified cells, skilled artisans will be able to select from known means to achieve a desired result. Means for introducing recombinant DNA into plant tissue include, but are not limited to, direct DNA uptake (Paszkowski, J. et al. (1984) EMBO J. 3:2717), electroporation (Fromm, M. et al. (1985) Proc. Natl. Acad. Sci. USA 82:5824), microinjection (Crossway, A. et al. (1986) Mol. Gen. Genet. 202:179), or T-DNA mediated transfer from *Agrobacterium tumefaciens* to the plant tissue. There appears to be no fundamental limitation of T-DNA transformation to the natural host range of Agrobacterium. Successful T-DNA-mediated transformation of monocots (Hooykaas-Van Slogteren, G. et al. (1984) Nature 311:763), gymnosperm (Dandekar, A. et al. (1987) Biotechnology 5:587) and algae (Ausich, R., EPO application 108,580) has been reported. Representative T-DNA vector systems are described in the following references: An, G. et al. (1985) EMBO J. 4:277; Herrera-Estrella, L. et al. (1983) Nature 303:209; Herrera-Estrella, L. et al. (1983) EMBO J. 2:987; Herrera-Estrella, L. et al. (1985) in *Plant Genetic Engineering*, New York: Cambridge University Press, p. 63. Once introduced into the plant tissue, the expression of the structural gene may be assayed by any means known to the art, and expression may be measured as mRNA transcribed or as protein synthesized. Techniques are known for the in vitro culture of plant tissue, and in a number of cases, for regeneration into whole plants. Procedures for transferring the introduced expression complex to commercially useful cultivars are known to those skilled in the art.

In one of its preferred embodiments the invention disclosed herein comprises expression in plant cells of a synthetic insecticidal structural gene under control of a plant expressible promoter, that is to say, by inserting the insecticide structural gene into T-DNA under control of a plant expressible promoter and introducing the T-DNA containing the insert into a plant cell using known means. Once plant cells expressing a synthetic insecticidal structural gene under control of a plant expressible promoter are obtained, plant tissues and whole plants can be regenerated therefrom using methods and techniques well-known in the art. The regenerated plants are then reproduced by conventional means and the introduced genes can be transferred to other strains and cultivars by conventional plant breeding techniques.

The introduction and expression of the synthetic structural gene for an insecticidal protein can be used to protect a crop from infestation with common insect pests. Other uses of the invention, exploiting the properties of other insecticide structural genes introduced into other plant species will be readily apparent to those skilled in the art. The invention in principle applies to. introduction of any synthetic insecticide structural gene into any plant species into which foreign DNA (in the preferred embodiment T-DNA) can be introduced and in which said DNA can remain stably replicated. In general, these taxa presently include, but are not limited to, gymnosperms and dicotyledonous plants, such as sunflower (family Compositeae), tobacco (family Solanaceae), alfalfa, soybeans and other legumes (family Leguminoseae), cotton (family Malvaceae), and most vegetables, as well as monocotyledonous plants. A plant containing in its tissues increased levels of insecticidal protein will control less susceptible types of insect, thus providing advantage over present insecticidal uses of *Bt*. By incorporation of the insecticidal protein into the tissues of a plant, the present invention additionally provides advantage over present uses of insecticides by eliminating instances of nonuniform application and the costs of buying and applying insecticidal preparations to a field. Also, the present invention eliminates the need for careful timing of application of such preparations since small larvae are most sensitive to insecticidal protein and the protein is always present, minimizing crop damage that would otherwise result from preapplication larval foraging.

This invention combines the specific teachings of the present disclosure with a variety of techniques and expedients known in the art. The choice of expedients depends on variables such as the choice of insecticidal protein from a *Bt* strain, the extent of modification in preferred codon usage, manipulation of sequences considered to be destabilizing to RNA or sequences prematurely terminating transcription, insertions of restriction sites within the design of the synthetic gene to allow future nucleotide modifications, addition of introns or enhancer sequences to the 5' and/or 3' ends of the synthetic structural gene, the promoter region, the host in which a promoter region/structural gene combination is expressed, and the like. As novel insecticidal proteins and toxic polypeptides are discovered, and as sequences responsible for enhanced cross-expression (expression of a foreign structural gene in a given host) are elucidated, those of ordinary skill will be able to select among those elements to produce "improved" synthetic genes for desired-proteins having agronomic value. The fundamental aspect of the present invention is the ability to synthesize a novel gene coding for an insecticidal protein, designed so that the protein will be expressed at an enhanced level in plants, yet so that it will retain its inherent property of insect toxicity and retain or increase its specific insecticidal activity.

EXAMPLES

The following Examples are presented as illustrations. of embodiments of the present invention. They do not limit the scope of this invention, which is determined by the claims.

The following strains were deposited with the Patent Culture Collection, Northern Regional Research Center, 1815 N. University Street, Peoria, Ill. 61604.

Strain	Deposited on	Accession #
E. coli MC1061 (p544-HindIII)	6 October 1987	NRRL B-18257
E. coli MC1061 (p544Pst-Met5)	6 October 1987	NRRL B-18258

The deposited strains are provided for the convenience of those in the art, and are not necessary to practice the present invention, which may be practiced with the present disclosure in combination with publicly available protocols, information, and materials. *E. coli* MC1061, a good host for plasmid transformations, was disclosed by Casadaban, M. J. and Cohen, S. N. (1980) J. Mol. Biol. 138:179–207.

Example 1

Design of the Synthetic Insecticidal Crystal Protein Gene

(i) Preparation of Toxic Subclones of the *Btt* Gene

Construction, isolation, and characterization of pNSB544 is disclosed by Sekar, V. et al. (1987) Proc. Natl. Acad. Sci. USA 84:7036–7040, and Sekar, V. and Adang, M. J., U.S. patent application Ser. No. 108,285, filed Oct. 13, 1987, which is hereby incorporated by reference. A 3.0 kbp HindIII fragment carrying the crystal protein gene of pNSBP544 is inserted into the HindIII site of pIC-2OH (Marsh, J. L. et al. (1984) Gene 32:481–485), thereby yielding a plasmid designated p544-HindIII, which is on deposit. Expression in *E. coli* yields a 73 kDa crystal protein in addition to the 65 kDa species characteristic of the crystal protein obtained from *Btt* isolates.

A 5.9 kbp BamHI fragment carrying the crystal protein gene is removed from pNSBP544 and inserted into BamHI-linearized pIC-2OH DNA. The resulting plasmid, p405/44-7, is digested with BglII and religated, thereby removing Bacillus sequences flanking the 3'-end of the crystal protein gene. The resulting plasmid, p405/54-12, is digested with PstI and religated, thereby removing Bacillus sequences flanking the 5'-end of the crystal protein and about 150 bp from the 5'-end of the crystal protein structural gene. The resulting plasmid, p405/81-4, is digested with SphI and PstI and is mixed with and ligated to a synthetic linker having the following structure:

```
          SD          MetThrAla      (SEQ ID NO:5)
     5'CAGGATCCAACAATGACTGCA3'
3'GTACGTCCTAGGTTGTTACTG5'
  SphI                    PstI
```

(SD indicates the location of a Shine-Dalgarno prokaryotic ribosome binding site.) The resulting plasmid, p544Pst-Met5, contains a structural gene encoding a protein identical to one encoded by pNSBP544 except for a deletion of the amino-terminal 47 amino acid residues. The nucleotide sequence of the *Btt* coding region in p544Pst-Met5 is presented in FIG. 1. In bioassays (Sekar and Adang, U.S. patent application Ser. No. 108,285, supra), the proteins encoded by the full-length *Btt* gene in pNSBP544 and the N-terminal deletion derivative, p544Pst-Met5, were shown to be equally toxic. All of the plasmids mentioned above have their crystal protein genes in the same orientation as the lacZ gene of the vector.

(ii) Modification of Preferred Codon Usage

Table 1 presents the frequency of codon usage for (A) dicot proteins, (B) *Bt* proteins, (C) the synthetic *Btt* gene, and (D) monocot proteins. Although some codons for a particular amino acid are utilized to approximately the same extent by both dicot and *Bt* proteins (e.g., the codons for serine), for the most part, the distribution of codon frequency varies significantly between dicot and *Bt* proteins, as illustrated in columns A and B in Table 1.

TABLE 1

Frequency of Codon Usage

		Distribution Fraction			
Amino Acid	Codon	(A)Dicot Genes	(B)Bt Genes	(C)Synthetic Btt Gene	(D)Monocot Genes
Gly	GGG	0.12	0.08	0.13	0.21
Gly	GGA	0.38	0.53	0.37	0.17
Gly	GGT	0.33	0.24	0.34	0.18
Gly	GGC	0.16	0.16	0.16	0.43
Glu	GAG	0.51	0.13	0.52	0.75
Glu	GAA	0.49	0.87	0.48	0.25
Asp	GAT	0.58	0.68	0.56	0.27
Asp	GAC	0.42	0.32	0.44	0.73
Val	GTG	0.29	0.15	0.30	0.36
Val	GTA	0.12	0.32	0.10	0.08
Val	GTT	0.39	0.29	0.35	0.19
Val	GTC	0.20	0.24	0.25	0.37
Ala	GCG	0.06	0.12	0.06	0.22
Ala	GCA	0.25	0.50	0.24	0.16
Ala	GCT	0.42	0.32	0.41	0.24
Ala	GCC	0.27	0.06	0.29	0.38
Lys	AAG	0.61	0.13	0.58	0.86
Lys	AAA	0.39	0.87	0.42	0.14
Asn	AAT	0.45	0.79	0.44	0.25
Asn	AAC	0.55	0.21	0.56	0.75
Met	ATG	1.00	1.00	1.00	1.00
Ile	ATA	0.18	0.30	0.20	0.11
Ile	ATT	0.45	0.57	0.43	0.24
Ile	ATC	0.37	0.13	0.37	0.64
Thr	ACG	0.08	0.14	0.07	0.20
Thr	ACA	0.27	0.68	0.27	0.14
Thr	ACT	0.35	0.14	0.34	0.19

TABLE 1-continued

		Frequency of Codon Usage			
Thr	ACC	0.30	0.05	0.32	0.46
Trp	TGG	1.00	1.00	1.00	1.00
End	TGA	0.33	0.00	0.00	0.34
Cys	TGT	0.44	0.33	0.33	0.30
Cys	TGC	0.56	0.67	0.67	0.70
End	TAG	0.19	0.00	0.00	0.36
End	TAA	0.48	1.00	1.00	0.30
Tyr	TAT	0.43	0.81	0.43	0.21
Tyr	TAC	0.57	0.19	0.57	0.79
Phe	TTT	0.45	0.75	0.44	0.25
Phe	TTC	0.55	0.25	0.56	0.75
Ser	AGT	0.14	0.25	0.13	0.08
Ser	AGC	0.18	0.13	0.19	0.26
Ser	TCG	0.06	0.08	0.06	0.14
Ser	TCA	0.19	0.19	0.17	0.11
Ser	TCT	0.25	0.25	0.27	0.15
Ser	TCC	0.18	0.10	0.17	0.25
Arg	AGG	0.25	0.09	0.23	0.26
Arg	AGA	0.30	0.50	0.32	0.09
Arg	CGG	0.04	0.14	0.05	0.13
Arg	CGA	0.08	0.14	0.09	0.04
Arg	CGT	0.21	0.09	0.23	0.12
Arg	CGC	0.11	0.05	0.09	0.36
Gln	CAG	0.41	0.18	0.39	0.46
Gln	CAA	0.59	0.82	0.61	0.54
His	CAT	0.54	0.90	0.50	0.33
His	CAC	0.46	0.10	0.50	0.67
Leu	TTG	0.26	0.08	0.27	0.14
Leu	TTA	0.10	0.46	0.12	0.03
Leu	CTG	0.09	0.04	0.10	0.28
Leu	CTA	0.08	0.21	0.10	0.10
Leu	CTT	0.28	0.15	0.18	0.15
Leu	CTC	0.19	0.06	0.22	0.31
Pro	CCG	0.09	0.20	0.08	0.23
Pro	CCA	0.42	0.56	0.44	0.34
Pro	CCT	0.32	0.24	0.32	0.17
Pro	CCC	0.17	0.00	0.16	0.26

154 coding sequences of dicot nuclear genes were used to compile the codon usage table. The pooled dicot coding sequences, obtained from Genbank (release 55) or, when no Genbank file name is specified, directly from the published source, were:

GENUS/SPECIES	GENBANK	PROTEIN	REF
Antirrhinum majus	AMACHS	Chalcone synthetase	
Arabidopsis thaliana	ATHADH	Alcohol dehydrogenase	
	ATHH3GA	Histone 3 gene 1	
	ATHH3GB	Histone 3 gene 2	
	ATHH4GA	Histone 4 gene 1	
	ATHLHCP1	CAB	
	ATHTUBA	α tubulin	
		5-enolpyruvyl4hifate 3-phosphate synthetase	1
Bertholletia excelsa		High methionine storage protein	2
Brassica campestris		Acyl carrier protein	3
Brassica napus	BNANAP	Napin	
Brassica oleacea	BOLSLSGR	S-locus specific glycoprotein	
Canavalia ensiformis	CENCONA	Concanavalin A	
Carica papaya	CPAPAP	Papain	
Chlamdomonas reinhardtii	CREC552	Preapocytochrome	
	CRERBCS1	RuBPC small subunit gene 1	
	CRERBCS2	RuBPC small subunit gene 2	
Cucurbita pepo	CUCPHT	Phytochrome	
Cucumis sativus	CUSGMS	Glyoxosomal malate synthetase	
	CUSLHCPA	CAB	
	CUSSSU	RuBPC small subunit	

TABLE 1-continued

	Frequency of Codon Usage		
Daucus carota	DAREXT	Extensin	
	DAREXTR	33 kD extensin related protein	
Dolichos biflorus	DBILECS	seed lectin	
Flaveria trinervia	FTRBCR	RuBPC small subunit	
Glycine max	SOY7SAA	75 storage protein	
	SOYACT1G	Actin 1	
	SOYCIIPI	CII protease inhibitor	
	SOYGLYA1A	Glycinin A1a Bx subunits	
	SOYGLYAAB	Glycinin A5A4B3 subunits	
	SOYGLYAB	Glycinin A3/b4 subunits	
	SOYGLYR	Glycinin A2B1a subunits	
	SOYHSP175	Low M W heat shock proteins	
	SOYLGBI	Leghemoglobin	
	SOYLEA	Lectin	
	SOYOX	Lipoxygenase 1	
	SOYNOD20G	20 kDa nodulin	
	SOYNOD23G	23 kDa nodulin	
	SOYNOD24H	24 kDa nodulin	
	SOYNOD26B	26 kDa nodulin	
	SOYNOD26R	26 kDa nodulin	
	SOYNOD27R	27 kDa nodulin	
	SOYNOD35M	35 kDa nodulin	
	SOYNOD75	75 kDa nodulin	
	SOYNODR1	Nodulin C51	
	SOYNODR2	Nodulin E27	
	SOYPRP1	Proline rich protein	
	SOYRUBP	RuBPC small subunit	
	SOYURA	Urease	
	SOYHSP26A	Heat shock protein 26A	
		Nuclear-encoded chloroplast heat shock protein	4
		22 kDa nodulin	5
		β1 tubulin	6
		β2 tubulin	6
Gossypium hirsutum		Seed α globulin (vicilin)	7
		Seed β globulin (vicilin)	7
Helianthus annus	HNNRUBCS	RuBPC small subunit	
		2S albumin seed storage protein	8
Ipomoea batatas		Wound-induced catalase	9
Lemna gibba	LGIAB19	CAB	
	LGIR5BPC	RuBPC small subunit	
Lupinus luteus	LUPLBR	leghemoglobin I	
Lycopersicon esculentum	TOMBIOBR	Biotin binding protein	
	TOMETHYBR	Ethylene biosynthesis protein	
	TOMPG2AR	Polygalacturonase-2a	
	TOMPSI	Tomato photosystem I protein	
	TOMRBCSA	RuBPC small subunit	
	TOMRBCSB	RuBPC small subunit	
	TOMRBCSC	RuBPC small subunit	
	TOMRBCSD	RuBPC small subunit	
	TOMRRD	Ripening related protein	
	TOMWIPIG	Wound induced proteinase inhibitor I	
	TOMWIPII	Wound induced proteinase inhibitor II	
		CAB 1A	10
		CAB 1B	10
		CAB 3C	10
		CAB 4	11
		CAB 5	11
Medicago sativa	ALFLB3R	Leghemoglobin III	
Mesembryanthemum crystallinum		RuBPC small subunit	12

TABLE 1-continued

		Frequency of Codon Usage	
Nicotiana plumbaginifolia	TOBATP21	Mitochondrial ATP synthase β subunit	
		Nitrate reductase	13
		Glutamine synthetase	14
Nicotiana tabacum	TOBECH	Endochitinase	
	TOBGAPA	A subunit of chloroplast G3PD	
	TOBGAPB	B subunit of chloroplast G3PD	
	TOBGAPC	C subunit of chloroplast G3PD	
	TOBPR1AR	Pathogenesis related protein 1a	
	TOBPR1CR	Pathogenesis-related protein 1c	
	TOBPRPR	Pathogenesis related protein 1b	
	TOBPXDLF	Peroxidase	
	TOBRBPCO	RuBPC small subunit	
	TOBTHAUR	TMV-induced protein homologous to thaumatin	
Perseus americana	AVOCEL	Cellulase	
Petroselinum hortense	PHOCHL	Chalcone synthase	
Petunia sp.	PETCAB13	CAB 13	
	PETCAB22L	CAB 22L	
	PETCAB22R	CAB 22R	
	PETCAB25	CAB 25	
	PETCAB37	CAB 37	
	PETCAB91R	CAB 91R	
	PETCHSR	Chalcone synthase	
	PETGCR1	Glycine-rich protein	
	PETRBCS08	RuBPC small subunit	
	PETRBCS11	RuBPC small subunit	
		70 kDa heat shock protein	15
Phascolus vulgaris	PHVCHM	Chitinase	
	PHVDLECA	Phytohemogglutinin E	
	PHVDLECB	Phytohemagglutinin L	
	PHVGSR1	Glutamine synthetase 1	
	PHVGSR2	Glutamine synthetase 2	
	PHVLBA	Leghemogobin	
	PHVLECT	Lectin	
	PHVPAL	Phenylalanine ammonia lyase	
	PHVPHASAR	α phaseolin	
	PHVPHASBR	β phaseolin	
		Arcelin seed protein	16
		Chalcone synthase	17
Pisum sativum	PEAALB2	Seed albumin	
	PEACAB80	CAB	
	PEAGSR1	Glutamine synthetase (nodule)	
	PEALECA	Lectin	
	PEALEGA	Legumin	
	PEARUBPS	RuBPC small subunit	
	PEAVIC2	Vicilin	
	PEAVIC4	Vicilin	
	PEAVIC7	Vicilin	
		Alcohol dehydrogenase 1	18
		Glutamine synthetase (leaf)	19
		Glutamine synthetase (root)	19
		Histone 1	20
		Nuclear encoded chloroplast heat shock protein	4
Raphanus sativus		RuBPC small subunit	21
Ricinus communis	RCCAGG	Agglutinin	
	RCCRICIN	Ricin	
	RCCICL4	Isocitrate lyase	
Silene pratensis	SIPFDX	Ferrodoxin precursor	
Solanum tuberosum	POTPAT	Patatin	
	POTINHWI	Wound-induced proteinase inhibitor	
	POTLS1G	Light-inducible tissue specific ST-LS1 gene	
	POTI2G	Wound-induced proteinase inhibitor II	
	POTRBCS	RuBPC small subunit	
		Sucrose synthetase	22
Spinacia oleracea	SPIACPI	Acyl carrier protein 1	
	SPIOEC16	16 kDa photosynthetic oxygen-evolving protein	
	SPIOEC23	23 kDa photosynthetic oxygen-evolving protein	
	SPIPCG	Plastocyanin	
	SPIPS33	33 kDa photosynthetic water oxidation complex precursor	
		Glycolate oxidase	23
Vicia faba	VFALBA	Leghemoglobin	
	VFALEB4	Legumin B	
		Vicillin	24

Pooled 53 monocot coding sequences obtained from Genbank (release 55) or, when no Genbank file name is specified, directly from the published source, were:

GENUS/SPECIES	GENBANK	PROTEIN	REF
Avena sallva	ASTAP3R	Phytochrome 3	
Hordeum vulgare	BLYALR	Aleurain	
	BLYAMY1	α amylase 1	
	BLYAMY2	α amylase 2	
	BLYCHORD1	Hordein C	
	BLYGLUCB	β glucanase	
	BLYHORB	B1 hordein	
	BLYPAPI	Amylase/protease inhibitor	
	BLYTH1AR	Toxin α hordothionin	
	BLYUBIQR	Ubiquitin	
		Histone 3	25
		Leaf specific thionin 1	26
		Leaf specific thionin 2	26
		Plastocyanin	27
Oryta sativa	RICGLUTG	Glutelin	
		Glutelin	28
Triticum aestivum	WHTAMYA	α amylase	
	WHTCAB	CAB	
	WHTEMR	Em protein	
	WHTGIR	gibberellin responsive protein	
	WHTGLGB	γ gliadin	
	WHTGLIABA	α/β gliadin Class AII	
	WHTGLUT1	High MW glutenin	
	WHTH3	Histone 3	
	WHTH4091	Histone 4	
	WHTRBCB	RuBPC small subunit	
Secale cereale	RYESECGSR	γ secalin	
Zea mays	MZEA1G	40.1 kD A1 protein (NADPH-dependent reductase)	
	MZEACT1G	Actin	
	MZEADH11F	Alcohol dehydrogenase 1	
	MZEADH2NR	Alcohol dehydrogenase 2	
	MZEALD	Aldolase	
	MZEANT	ATP/ADP translocator	
	MZEEG2R	Glutelin 2	
	MZEGGST3B	Glutathione S transferase	
	MZEH3C2	Histone 3	
	MZEH4C14	Histone 4	
	MZEHSP701	70 kD Heat shock protein, exon 1	

TABLE 1-continued

Frequency of Codon Usage		
MZEHSP702	70 kD Heat shock protein, exon 2	
MZELHCP	CAB	
MZEMPL3	Lipid body surface protein L3	
MZEPEPCR	Phosphoenolyruvate carboxylase	
MZERBCS	RuBPC small subunit	
MZESUSYSG	Sucrose synthetase	
MZETPI2	Triosephosphate isomerase 1	
MZEZEA20M	19 kD zein	
MZEZEA30M	19 kD zein	
MZEZE15A3	15 kD zein	
MZEZE16	16 kD zein	
MZEZE19A	19 kD zein	
MZEZE22A	22 kD zein	
MZEZE22B	22 kD zein	
	Catalase 2	29
	Regulatory C1 locus	30

Bt codons were obtained from analysis of coding sequences of the following genes:
Bt var. kurstaki HD-73, 6.6 kb HindIII fragment (Kronstad et al. (1983) J. Bacteriol. 154:419–428);
Bt var. kurstaki HD-1, 5.3 kb fragment (Adang et al. (1987) in Biotechnology in Invertebrate Pathology and Cell Culture, K. Maramorosh (ed.), Academic Press, Inc. New York, pp. 85–99);
Bt var. kurstaki HD-1, 4.5 kb fragment (Schnepf and Whiteley (1985) J. Biol. Chem. 260:6273–6280); and
Bt var. tenebrionis, 3.0 kb HindIII fragment (Sekar et al. (1987) Proc. Natl. Acad. Sci. 84:7036–7040).

REFERENCES
1. Klee, H. J. et al. (1987) Mol. Gen. Genet. 210:437–442.
2. Altenbach, S. B. et al. (1987) Plant Mol. Biol. 8:239–250.
3. Rose, R. E. et al. (1987) Nucl. Acids Res. 15:7197.
4. Vierling, E. et al. (1988) EMBO J. 7:575–581.
5. Sandal, N. N. et al. (1987) Nucl. Acids Res. 15:1507–1519.
6. Tingey, S. V. et al. (1987) EMBO J. 6:1–9.
7. Chlan, C. A. et al. (1987) Plant Mol. Biol. 9:533–546.
8. Allen, R. D. et al. (1987) Mol. Gen. Genet. 210:211–218.
9. Sakajo, S. et al. (1987) Eur. J. Biochem. 165:437–442.
10. Pirersky, E. et al. (1987) Plant Mol. Biol. 9:109–120.
11. Ray, J. et al. (1987) Nucl. Acids Res. 15:10587.
12. DeRocjer, E. J. et al. (1987) Nucl. Acids Res. 15:6301.
13. Calza, R. et al. (1987) Mol. Gen. Genet. 209:552–562.
14. Tingey, S. V. and Coruzzi, G. M. (1987) Plant Phys. 84:366–373.
15. Winter, J. et al. (1988) Mol. Gen. Genet. 211:315–319.
16. Osborn, T. C. et al. (1988) Science 240:207–210.
17. Ryder, T. B. et al. (1987) Mol. Gen. Genet. 210:219–233.
18. Llewellyn, D. J. et al. (1987) J. Mol. Biol. 195:115–123.
19. Tingey, S. V. et al. (1987) EMBO J. 6:1–9.
20. Gantt, J. S. and Key, J. L. (1987) Eur. J. Biochem. 166:119–125.
21. Guidet, F. and Fourcroy, P. (1988) Nucl. Acids Res. 16:2336.
22. Salanoubat, M. and Belliard, G. (1987) Gene 60:47–56.
23. Volokita, M. and Somerville, C. R. (1987) J. Biol. Chem. 262:15825–15828.
24. Bassner, R. et al. (1987) Nucl. Acids Res. 15:9609.
25. Chojecki, J. (1986) Carlsberg Res. Commun. 51:211–217.
26. Bohlmann, H. and Apel, K. (1987) Mol. Gen. Genet. 207:446–454.
27. Nielsen, P. S. and Gausing, K. (1987) FEBS Lett. 225:159–162.
28. Higuchi, W. and Fukazawa, C. (1987) Gene 55:245–253.
29. Bethards, L. A. et al. (1987) Proc. Natl. Sci. USA 84:6830–6834.
30. Paz-Ares, J. et al. (1987) EMBO J. 6:3553–3558.

For example, dicots utilize the AAG codon for lysine with a frequency of 61% and the AAA codon with a frequency of 39%. In contrast, in *Bt* proteins the lysine codons AAG and AAA are used with a frequency of 13% and 87%, respectively. It is known in the art that seldom used codons are generally detrimental to that system and must be avoided or used judiciously. Thus, in designing a synthetic gene encoding the *Btt* crystal protein, individual amino acid codons found in the original *Btt* gene are altered to reflect the codons preferred by dicot genes for a particular amino acid. However, attention is given to maintaining the overall distribution of codons for each amino acid within the coding region of the gene. For example, in the case of alanine, it can be seen from Table 1 that the codon GCA is used in *Bt* proteins with a frequency of 50%, whereas the codon GCT is the preferred codon in dicot proteins. In designing the synthetic *Btt* gene, not all codons for alanine in the original *Bt* gene are replaced by GCT; instead, only some alanine codons are changed to GCT while others are replaced with different alanine codons in an attempt to preserve the overall distribution of codons for alanine used in dicot proteins. Column C in Table 1 documents that this goal is achieved; the frequency of codon usage in dicot proteins (column A) corresponds very closely to that used in the synthetic *Btt* gene (column C).

In similar manner, a synthetic gene coding for insecticidal crystal protein can be optimized for enhanced expression in monocot plants. In Table 1, column D, is presented the frequency of codon usage of highly expressed monocot proteins.

Because of the degenerate nature of the genetic code, only part of the variation contained in a gene is expressed in this protein. It is clear that variation between degenerate base frequencies is not a neutral phenomenon since systematic codon preferences have been reported for bacterial, yeast and mammalian genes. Analysis of a large group of plant gene sequences indicates that synonymous codons are used differently by monocots and dicots. These patterns are also distinct from those reported for *E. coli*, yeast and man.

In general, the plant codon usage pattern more closely resembles that of man and other higher eukaryotes than unicellular organisms, due to the overall preference for G+C content in codon position III. Monocots in this sample share the most commonly used codon for 13 of 18 amino acids as that reported for a sample of human genes (Grantham et al. (1986 supra), although dicots favor the most commonly used human codon in only 7 of 18 amino acids.

Discussions of plant codon usage have focused on the differences between codon choice in plant nuclear genes and in chloroplasts. Chloroplasts differ from higher plants in that they encode only 30 tRNA species. Since chloroplasts have restricted their tRNA genes, the use of preferred codons by chloroplast-encoded proteins appears more extreme. However, a positive correlation has been reported between the level of isoaccepting tRNA for a given amino acid and the frequency with which this codon is used in the chloroplast genome (Pfitzinger et al. (1987) Nucl. Acids Res. 15:1377–1386).

Our analysis of the plant genes sample confirms earlier reports that the nuclear and chloroplast genomes in plants have distinct coding strategies. The codon usage of monocots in this sample is distinct from chloroplast usage, sharing the most commonly used codon for only 1 of 18 amino acids. Dicots in this sample share the most commonly used codon of chloroplasts in only 4 of 18 amino acids. In general, the chloroplast codon profile more closely resembles that of unicellular organisms, with a strong bias towards the use of A+T in the degenerate third base.

In unicellular organisms, highly expressed genes use a smaller subset of codons than do weakly expressed genes although the codons preferred are distinct in some cases. Sharp and Li (1986) Nucl. Acids Res. 14:7734–7749 report that codon usage in 165 *E. coli* genes reveals a positive correlation between high expression and increased codon bias. Bennetzen and Hall (1982) supra have described a similar trend in codon selection in yeast. Codon usage in these highly expressed genes correlates with the abundance of isoaccepting tRNAs in both yeast and *E coli*. It has been proposed that the good fit of abundant yeast and *E. coli*

mRNA codon usage to isoacceptor tRNA abundance promotes high translation levels and high steady state levels of these proteins. This strongly suggests that the potential for high levels of expression of plant genes in yeast or *E. coli* is limited by their codon usage. Hoekema et al. (1987) supra report that replacement of the 25 most favored yeast codons with rare codons in the 5' end of the highly expressed gene PGK1 leads to a decrease in both mRNA and protein. These results indicate that codon bias should be emphasized when engineering high expression of foreign genes in yeast and other systems.

(iii) Sequences within the *Btt* Coding Region having Potentially Destabilizing Influences

Analysis of the *Btt* gene reveals that the A+T content represents 64% of the DNA base composition of the coding region. This level of A+T is about 10% higher than that found in a typical plant coding region. Most often, high A+T regions are found in intergenic regions. Also, many plant regulatory sequences are observed to be AT-rich. These observations lead to the consideration that an elevated A+T content within the *Btt* coding region may be contributing to a low expression level in plants. Consequently, in designing a synthetic *Btt* gene, the A+T content is decreased to more closely approximate the A+T levels found in plant proteins. As illustrated in Table 3, the A+T content is lowered to a level in keeping with that found in coding regions of plant nuclear genes. The synthetic *Btt* gene of this invention has an A+T content of 55%.

TABLE 3

Adenine + Thymine Content in Btt Coding Region

	Base					
Coding Region	G	A	T	C	%G+C	%A+T
Natural Btt gene	341	633	514	306	36	64
Synthetic Btt gene	392	530	483	428	45	55

In addition, the natural *Btt* gene is scanned for sequences that are potentially destabilizing to *Btt* RNA. These sequences, when identified in the original *Btt* gene, are eliminated through modification of nucleotide sequences. Included in this group of potentially destabilizing sequences are:

(a) plant polyadenylation signals (as described by Joshi (1987) Nucl. Acids Res. 15:9627–9640). In eukaryotes, the primary transcripts of nuclear genes are extensively processed (steps including 5'-capping, intron splicing, polyadenylation) to form mature and translatable mRNAs. In higher plants, polyadenylation involves endocucleolytic cleavage at the polyA site followed by the addition of several A residues to the cleaved end. The selection of the polyA site is presumed to be cis-regulated. During expression of *Bt* protein and RNA in different plants, the present inventors have observed that the polyadenylated mRNA isolated from these expression systems is not full-length but instead is truncated or degraded. Hence, in the present invention it was decided to minimize possible destabilization of RNA through elimination of potential polyadenylation signals within the coding region of the synthetic *Btt* gene. Plant polyadenylation signals including AATAAA, AATGAA, AATAAT, AATATT, GATAAA, GATAAA, and AATAAG motifs do not appear in the synthetic *Btt* gene when scanned for 0 mismatches of the sequences.

(b) polymerase II termination sequence, CAN_{7-9} AGTNNAA. This sequence was shown (Vankan and Filipowicz (1988) EMBO J. 7:791–799) to be next to the 3' end of the coding region of the U2 snRNA genes of *Arabidopsis thaliana* and is believed to be important for transcription termination upon 3' end processing. The synthetic *Btt* gene is devoid of this termination sequence.

(c) CUUCGG hairpins, responsible for extraordinarily stable RNA secondary structures associated with various biochemical processes (Tuerk et al. (1988) Proc. Natl. Acad. Sci. 85:1364–1368). The exceptional stability of CUUCGG hairpins suggests that they have an unusual structure and may function in organizing the proper folding of complex RNA structures. CUUCGG hairpin sequences are not found with either 0 or 1 mismatches in the *Btt* coding region.

(d) plant consensus splice sites, 5'=AAG:GTAAGT(SEQ ID NO:6) and 3'=TTTT(Pu)TTT(Pu)T(Pu)T(Pu)T(Pu)TGCAG:C(SEQ ID NO:7), as described by Brown et al. (1986) EMBO J. 5:2749–2758. Consensus sequences for the 5' and 3' splice junctions have been derived from 20 and 30 plant intron sequences, respectively. Although it is not likely that such potential splice sequences are present in *Bt* genes, a search was initiated for sequences resembling plant consensus splice sites in the synthetic *Btt* gene. For the 5' splice site, the closest match was with three mismatches. This gave 12 sequences of which two had G:GT. Only position 948 was changed because 1323 has the KpnI site needed for reconstruction. The 3'-splice site is not found in the synthetic *Btt* gene.

Thus, by highlighting potential RNA-destabilizing sequences, the synthetic *Btt* gene is designed to eliminate known eukaryotic regulatory sequences that affect RNA synthesis and processing.

Example 2

Chemical Synthesis of a Modified *Btt* Structural Gene

(i) Synthesis Strategy

The general plant for synthesizing linear double-stranded DNA sequences coding for the crystal protein from *Btt* is schematically simplified in FIG. 2. The optimized DNA coding sequence (FIGS. 1A–C) is divided into thirteen segments (segments A–M) to be synthesized individually, isolated and purified. As shown in FIG. 2, the general strategy begins by enzymatically joining segments A and M to form segments AM to which is added segment BL to form segment ABLM. Segment CK is then added enzymatically to make segment ABCKLM which is enlarged through addition of segments DJ, EI and RFH sequentially to give finally the total segment ABCDEFGHIJKLM, representing the entire coding region of the *Btt* gene.

FIG. 3 outlines in more detail the strategy used in combining individual DNA segments in order to effect the synthesis of a gene having unique restriction sites integrated into a defined nucleotide sequence. Each of the thirteen segments (A to M) has unique restriction sites at both ends, allowing the segment to be strategically spliced into a growing DNA polymer. Also, unique sites are placed at each end of the gene to enable easy transfer from one vector to another.

The thirteen segments (A to N) used to construct the synthetic gene vary in size. oligonucleotide pairs of approximately 75 nucleotides each are used to construct larger segments having approximately 225 nucleotide pairs. FIG. 3 documents the number of base pairs contained within each

segment and specifies the unique restriction sites bordering each segment. Also, the overall strategy to incorporate specific segments at appropriate splice sites is detailed in FIG. 3.

(ii) Preparation of Oligodeoxynucleotides

Preparation of oligodeoxynucleotides for use in the synthesis of a DNA sequence comprising a gene for *Btt* is carried out according to the general procedures described by Matteucci et al. (1981) J. Am. Chem. Soc. 103:3185–3192 and Beaucage et al. (1981) Tetrahedron Lett. 22:1859–1862. All oligonucleotides are prepared by the solid-phase phosphoramidite triester coupling approach, using an Applied Biosystems Model 380A DNA synthesizer. Deprotection and cleavage of the oligomers from the solid support are carried out according to standard procedures. Crude oligonucleotide mixtures are purified using an oligonucleotide purification cartridge (OTC, Applied Biosystems) as described by McBride et al. (1988) Biotechniques 6:362–367.

5'-phosphorylation of oligonucleotides is performed with T4 polynucleotide kinase. The reaction contains 2 μg oligonucleotide and 18.2 units polynucleotide kinase (Pharmacia) in linker kinase buffer (Maniatis (1982) *Cloning Manual*, Fritsch and Sambrook (eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). The reaction is incubated at 37° C. for 1 hour.

Oligonucleotides are annealed by first heating to 95° C. for 5 min. and then allowing complementary pairs to cool slowly to room temperature. Annealed pairs are reheated to 65° C., solutions are combined, cooled slowly to room temperature and kept on ice until used. The ligated mixture may be purified by electrophoresis through a 4% NuSieve agarose (FMC) gel. The band corresponding to the ligated duplex is excised, the DNA is extracted from the agarose and ethanol precipitated.

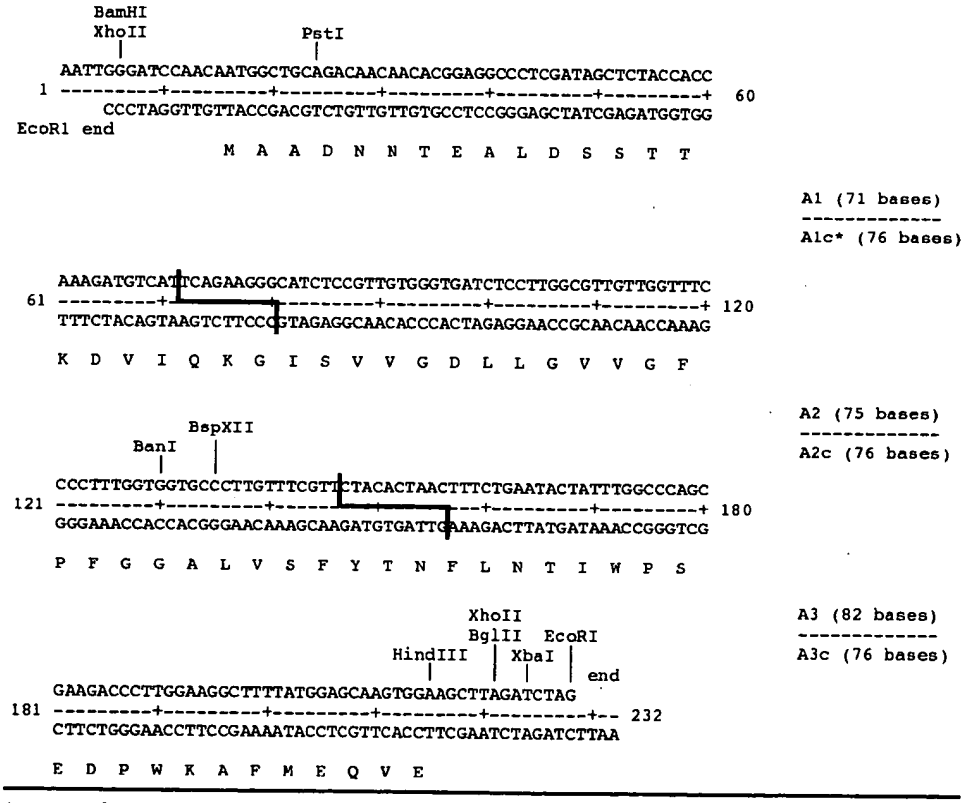
Ligations are carried out as exemplified by that used in M segment ligations. M segment DNA is brought to 65° C. for 25 min, the desired vector is added and the reaction mixture is incubated at 65° C. for 15 min. The reaction is slow cooled over 1½ hours to room temperature. ATP to 0.5 mM and 3.5 units of T4 DNA ligase salts are added and the reaction mixture is incubated for 2 hr at room temperature and then maintained overnight at 15° C. The next morning, vectors which had not been ligated to M block DNA were removed upon linearization by EcoRI digestion. Vectors ligated to the M segment DNA are used to transform *E. coli* MC1061. Colonies containing inserted blocks are identified by colony hybridization with ^{32}P-labelled oligonucleotide probes. The sequence of the DNA segment is confirmed by isolating plasmid DNA and sequencing using the dideoxy method of Sanger et al. (1977) Proc. Natl. Acad. Sci. 74:5463–5467.

(iii) Synthesis of Segment AM

Three oligonucleotide pairs (A1 and its complementary strand A1c, A2 and A2c and A3 and A3c) are assembled and ligated as described above to make up segment A. The nucleotide sequence of segment A is as follows:

TABLE 4

Nucleotide Sequence of Segment A



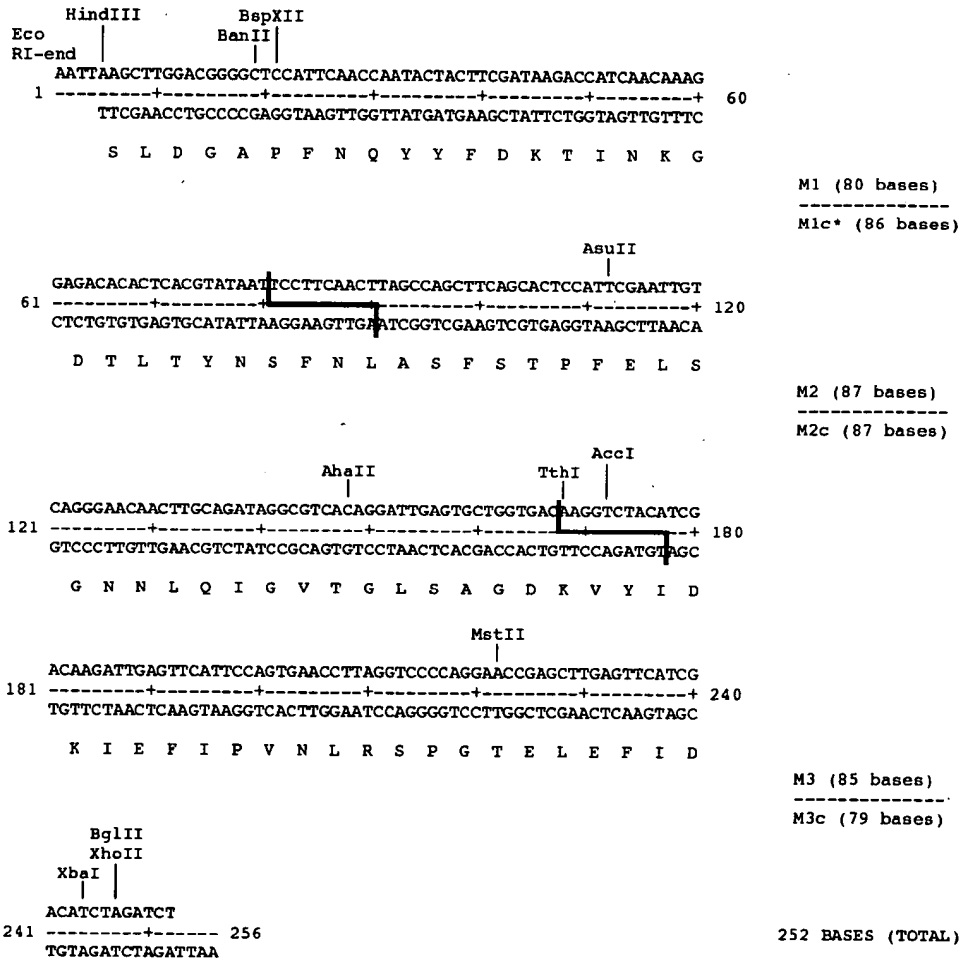
*c = complementary strand.

In Table 4, bold lines demarcate the individual oligonucleotides. Fragment A1 contains 71 bases, A1c has 76 bases, A2 has 75 bases, A2c has 76 bases, A3 has 82 bases and A3c has 76 bases. In all, segment A is composed of 228 base pairs and is contained between EcoRI restriction enzyme site and one destroyed EcoRI site (5')J. (Additional restriction sites within Segment A are indicated.) The EcoRI single-stranded cohesive ends allow segment A to be annealed and then ligated to the EcoRI-cut cloning vector, pIC20K.

Segment M comprises three oligonucleotide pairs: M1, 80 bases, M1c, 86 bases, M2, 87 bases, M2c, 87 bases, M3, 85 bases and M3c 79 bases. The individual oligonucleotides are annealed and ligated according to standard procedures as described above. The overall nucleotide sequence of segment M is:

TABLE 5

Nucleotide Sequence of Segment M



*c = complementary strand

In Table 5 bold lines demarcate the individual oligonucleotides. Segment M contains 252 base pairs and has destroyed EcoRI, restriction sites at both ends. (Additional restriction sites within segment M are indicated). Segment M is inserted into vector pIC20R at an EcoRI restriction site and cloned.

As proposed in FIG. 3, segment M is joined to segment A in the plasmid in which it is contained. Segment M is excised at the flanking restrictions sites from its cloning vector and spliced into pIC20K, harboring segment A, through successive digestions with HindIII followed by BglII. The pIC20K vector now comprises segment A joined to segment M with a HindIII site at the splice site (see FIG. 3). Plasmid pIC20K is derived from pIC20R by removing the ScaI-NdeI DNA fragment and inserting a HincII fragment containing an NPTI coding region. The resulting plasmid of 4.44 kb confers resistance to kanamycin on *E. coli.*

Example 3

Expression of Synthetic Crystal Protein Gene in Bacterial Systems

The synthetic *Btt* gene is designed so that it is expressed in the pIC20R-kan vector in which it is constructed. This expression is produced utilizing the initiation methionine of the lacZ protein of pIC20K. The wild-type *Btt* crystal protein sequence expressed in this manner has full insecticidal activity. In addition, the synthetic gene is designed to contain a BamHI site 5' proximal to the initiating methionine codon and a BglII site 3' to the terminal TAG translation stop codon. This facilitates the cloning of the insecticidal crystal

protein coding region into bacterial expression vectors such as pDR540 (Russell and Bennett, 1982). Plasmid pDR540 contains the TAC promoter which allows the production of proteins including *Btt* crystal protein under controlled conditions in amounts up to 10% of the total bacterial protein. This promoter functions in many gram-negative bacteria including *E. coli* and Pseudomonas.

Production of *Bt* insecticidal crystal protein from the synthetic gene in bacteria demonstrates that the protein produced has the expected toxicity to coleopteran insects. These recombinant bacterial strains in themselves have potential value as microbial insecticides, product of the synthetic gene.

Example 4

Expression of a Synthetic Crystal Protein Gene in Plants

The synthetic *Btt* crystal protein gene is designed to facilitate cloning into the expression cassettes. These utilize sites compatible with the BamHI and BglII restriction sites flanking the synthetic gene. Cassettes are available that utilize plant promoters including CaMV 35S, CaMV 19S and the ORF 24 promoter from T-DNA. These cassettes provide the recognition signals essential for expression of proteins in plants. These cassettes are utilized in the micro Ti plasmids such as pH575. Plasmids such as pH575 containing the synthetic *Btt* gene directed by plant expression signals are utilized in disarmed *Agrobacterium tumefaciens* to introduce the synthetic gene into plant genomic DNA. This system has been described previously by Adang et al. (1987) to express *Bt* var. *kurstaki* crystal protein gene in tobacco plants. These tobacco plants were toxic to feeding tobacco hornworms.

Example 5

Assay for Insecticidal Activity

Bioassays are conducted essentially as described by Sekar, V. et al. supra. Toxicity is assessed by an estimate of the LD_{50}. Plasmids are grown in *E. coli* JM105 (Yanisch-Perron, C. et al. (1985) Gene 33:103–119). On a molar basis, no significant differences in toxicity are observed between crystal proteins encoded by p544Pst-Met5, p544-HindIII, and pNSBP544. When expressed in plants under identical conditions, cells containing protein encoded by the synthetic gene are observed to be more toxic than those containing protein encoded by the native *Btt* gene. Immunoblots ("western" blots) of cell cultures indicated that those that are more toxic have more crystal protein antigen. Improved expression of the synthetic *Btt* gene relative to that of a natural *Btt* gene is seen as the ability to quantitate specific mRNA transcripts from expression of synthetic *Btt* genes on Northern blot assays.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 9

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1794 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATGACTGCAG ATAATAATAC GGAAGCACTA GATAGCTCTA CAACAAAAGA TGTCATTCAA     60

AAAGGCATTT CCGTAGTAGG TGATCTCCTA GGCGTAGTAG GTTTCCCGTT TGGTGGAGCG    120

CTTGTTTCGT TTTATACAAA CTTTTTAAAT ACTATTTGGC CAAGTGAAGA CCCGTGGAAG    180

GCTTTTATGG AACAAGTAGA AGCATTGATG GATCAGAAAA TAGCTGATTA TGCAAAAAAT    240

AAAGCTCTTG CAGAGTTACA GGGCCTTCAA AATAATGTCG AAGATTATGT GAGTGCATTG    300

AGTTCATGGC AAAAAAATCC TGTGAGTTCA CGAAATCCAC ATAGCCAGGG GCGGATAAGA    360

GAGCTGTTTT CTCAAGCAGA AAGTCATTTT CGTAATTCAA TGCCTTCGTT TGCAATTTCT    420

GGATACGAGG TTCTATTTCT AACAACATAT GCACAAGCTG CCAACACACA TTTATTTTTA    480

CTAAAAGACG CTCAAATTTA TGGAGAAGAA TGGGGATACG AAAAAGAAGA TATTGCTGAA    540

TTTTATAAAA GACAACTAAA ACTTACGCAA GAATATACTG ACCATTGTGT CAAATGGTAT    600

AATGTTGGAT TAGATAAATT AAGAGGTTCA TCTTATGAAT CTTGGGTAAA CTTTAACCGT    660

TATCGCAGAG AGATGACATT AACAGTATTA GATTTAATTG CACTATTTCC ATTGTATGAT    720
```

-continued

```
GTTCGGCTAT ACCCAAAAGA AGTTAAAACC GAATTAACAA GAGACGTTTT AACAGATCCA    780

ATTGTCGGAG TCAACAACCT TAGGGGCTAT GGAACAACCT TCTCTAATAT AGAAAATTAT    840

ATTCGAAAAC CACATCTATT TGACTATCTG CATAGAATTC AATTTCACAC GCGGTTCCAA    900

CCAGGATATT ATGGAAATGA CTCTTTCAAT TATTGGTCCG GTAATTATGT TTCAACTAGA    960

CCAAGCATAG GATCAAATGA TATAATCACA TCTCCATTCT ATGGAAATAA ATCCAGTGAA   1020

CCTGTACAAA ATTTAGAATT TAATGGAGAA AAAGTCTATA GAGCCGTAGC AAATACAAAT   1080

CTTGCGGTCT GGCCGTCCGC TGTATATTCA GGTGTTACAA AAGTGGAATT TAGCCAATAT   1140

AATGATCAAA CAGATGAAGC AAGTACACAA ACGTACGACT CAAAAAGAAA TGTTGGCGCG   1200

GTCAGCTGGG ATTCTATCGA TCAATTGCCT CCAGAAACAA CAGATGAACC TCTAGAAAAG   1260

GGATATAGCC ATCAACTCAA TTATGTAATG TGCTTTTTAA TGCAGGGTAG TAGAGGAACA   1320

ATCCCAGTGT TAACTTGGAC ACATAAAAGT GTAGACTTTT TTAACATGAT TGATTCGAAA   1380

AAAATTACAC AACTTCCGTT AGTAAAGGCA TATAAGTTAC AATCTGGTGC TTCCGTTGTC   1440

GCAGGTCCTA GGTTTACAGG AGGAGATATC ATTCAATGCA CAGAAAATGG AAGTGCGGCA   1500

ACTATTTACG TTACACCGGA TGTGTCGTAC TCTCAAAAAT ATCGAGCTAG AATTCATTAT   1560

GCTTCTACAT CTCAGATAAC ATTTACACTC AGTTTAGACG GGGCACCATT TAATCAATAC   1620

TATTTCGATA AAACGATAAA TAAAGGAGAC ACATTAACGT ATAATTCATT TAATTTAGCA   1680

AGTTTCAGCA CACCATTCGA ATTATCAGGG AATAACTTAC AAATAGGCGT CACAGGATTA   1740

AGTGCTGGAG ATAAAGTTTA TATAGACAAA ATTGAATTTA TTCCAGTGAA TTAA         1794
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 597 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160
```

-continued

```
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
    530                 535                 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575
```

-continued

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

Phe Ile Pro Val Asn
595

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1833 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGCTGCAG ACAACAACAC GGAGGCCCTC GATAGCTCTA CCACCAAAGA TGTCATTCAG 60

AAGGGCATCT CCGTTGTGGG TGATCTCCTT GGCGTTGTTG GTTTCCCCTT TGGTGGTGCC 120

CTTGTTTCGT TCTACACTAA CTTTCTGAAT ACTATTTGGC CCAGCGAAGA CCCTTGGAAG 180

GCTTTTATGG AGCAAGTGGA AGCTTTGATG GATCAGAAGA TCGCTGATTA TGCAAAGAAC 240

AAAGCTCTTG CTGAGCTCCA GGGCCTTCAG AACAACGTCG AAGATTATGT GAGTGCACTG 300

AGTTCATGGC AAAAGAATCC TGTGTCCTCA CGAAATCCAC ATAGCCAGGG GCGCATAAGG 360

GAGCTGTTCT CTCAAGCAGA AAGTCACTTC CGGAATTCAA TGCCTTCCTT TGCCATCTCT 420

GGGTACGAGG TTCTCTTTCT TACAACCTAC GCTCAAGCTG CCAACACACA TCTGTTCTTA 480

CTAAAAGACG CTCAAATCTA TGGTGAAGAA TGGGGATACG AGAAAGAAGA TATCGCTGAG 540

TTCTACAAGC GTCAACTAAA ACTTACTCAA GAGTATACTG ACCACTGTGT CAAATGGTAT 600

AATGTTGGAT TGGATAAGTT GAGAGGTTCA TCTTATGAAT CTTGGGTAAA CTTTAACCGG 660

TACCGCAGAG AGATGACATT GACAGTGCTC GACTTGATTG CACTATTTCC ATTGTATGAT 720

GTTCGACTCT ACCCAAAGGA GGTTAAAACC GAATTGACTA GAGACGTTTT AACCGATCCC 780

ATTGTCGGAG TCAACAACCT CAGAGGCTAC GGAACAACCT TCTCTAACAT AGAAAACTAC 840

ATTCGTAAAC CACATCTATT CGACTATCTG CACAGAATTC AGTTTCACAC GCGGTTCCAA 900

CCAGGATACT ATGGAAATGA CTCTTTCAAC TATTGGTCCG GTAATTATGT TTCAACTAGA 960

CCCAGCATAG GATCTAATGA CATCATCACC TCTCCATTCT ACGGAAACAA GTCCTCCGAG 1020

CCTGTGCAAA ACTTGGAGTT TAATGGAGAG AAAGTCTATA GAGCCGTGGC CAATACCAAT 1080

CTTGCCGTCT GGCCGTCCGC TGTGTACTCA GGTGTTACCA AAGTGGAATT CAGCCAATAC 1140

AATGATCAGA CAGATGAAGC AAGTACTCAA ACTTACGACT CAAAGAGGAA TGTTGGCGCG 1200

GTCAGCTGGG ATTCTATCGA TCAACTCCCT CCAGAAACCA CCGATGAACC TCTAGAGAAG 1260

GGTTATAGCC ATCAACTCAA TTACGTAATG TGCTTTCTCA TGCAGGGTAG TAGAGGTACC 1320

ATCCCAGTGT TAACTTGGAC TCACAAGAGT GTAGACTTCT TCAACATGAT TGATTCGAAA 1380

AAGATTACTC AACTTCCGTT GGTAAAGGCC TACAAGTTAC AATCTGGTGC TTCCGTTGTC 1440

GCAGGTCCTA GGTTTACAGG AGGAGATATC ATTCAATGCA CTGAGAATGG GTCCGCGGCA 1500

ACTATCTACG TTACACCTGA TGTGTCGTAC TCTCAAAAGT ATCGTGCTAG AATTCATTAT 1560

GCTTCTACCT CTCAGATAAC ATTCACACTA AGCTTGGACG GGGCTCCATT CAACCAATAC 1620

TACTTCGATA AGACCATCAA CAAAGGAGAC ACACTCACGT ATAATTCATT CAACTTAGCC 1680

AGCTTCAGCA CTCCATTCGA ATTGTCAGGG AACAACTTGC AGATAGGCGT CACAGGATTG 1740

AGTGCTGGTG ACAAGGTTTA CATCGACAAG ATTGAGTTCA TTCCAGTGAA CCTTAGGTCC 1800

-continued

CCAGGAACCG AGCTTGAGTT CATCGACATC TAG 1833

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 610 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

-continued

```
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
    530                 535                 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

Phe Ile Pro Val Asn Leu Arg Ser Pro Gly Thr Glu Leu Glu Phe Ile
        595                 600                 605

Asp Ile
    610
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAGGATCCAA CAATGAC 17

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAGGTAAGT 9

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TTTTUTTTUT UTUTUTGCAG C 21

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 456 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AATTGGGATC CAACAATGGC TGCAGACAAC AACACGGAGG CCCTCGATAG CTCTACCACC 60

CCCTAGGTTG TTACCGACGT CTGTTGTTGT GCCTCCGGGA GCTATCGAGA TGGTGGAAAG 120

ATGTCATTCA GAAGGGCATC TCCGTTGTGG GTGATCTCCT TGGCGTTGTT GGTTTCTTTC 180

TACAGTAAGT CTTCCCGTAG AGGCAACACC CACTAGAGGA ACCGCAACAA CCAAAGCCCT 240

TTGGTGGTGC CCTTGTTTCG TTCTACACTA ACTTTCTGAA TACTATTTGG CCCAGCGGGA 300

AACCACCACG GGAACAAAGC AAGATGTGAT TGAAAGACTT ATGATAAACC GGGTCGGAAG 360

ACCCTTGGAA GGCTTTTATG GAGCAAGTGG AAGCTTAGAT CTAGCTTCTG GGAACCTTCC 420

GAAAATACCT CGTTCACCTT CGAATCTAGA TCTTAA 456

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AATTAAGCTT GGACGGGGCT CCATTCAACC AATACTACTT CGATAAGACC ATCAACAAAG 60

TTCGAACCTG CCCCGAGGTA AGTTGGTTAT GATGAAGCTA TTCTGGTAGT TGTTTCGAGA 120

CACACTCACG TATAATTCCT TCAACTTAGC CAGCTTCAGC ACTCCATTCG AATTGTCTCT 180

GTGTGAGTGC ATATTAAGGA AGTTGAATCG GTCGAAGTCG TGAGGTAAGC TTAACACAGG 240

GAACAACTTG CAGATAGGCG TCACAGGATT GAGTGCTGGT GACAAGGTCT ACATCGGTCC 300

CTTGTTGAAC GTCTATCCGC AGTGTCCTAA CTCACGACCA CTGTTCCAGA TGTAGCACAA 360

GATTGAGTTC ATTCCAGTGA ACCTTAGGTC CCCAGGAACC GAGCTTGAGT TCATCGTGTT 420

CTAACTCAAG TAAGGTCACT TGGAATCCAG GGGTCCTTGG CTCGAACTCA AGTAGCACAT 480

CTAGATCTTG TAGATCTAGA TTAA 504

We claim:

1. A synthetic *Bacillus thuringiensis* (*B.t.*) gene which is expressed in descendant plant cells and encodes a pesticidal protein toxin, wherein said synthetic *B.t.* gene is produced by the process of:

selecting a *B.t.* pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native *B.t.* gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native *B.t.* coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native *B.t.* coding sequence;

obtaining a synthetic *B.t.* gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic *B.t.* gene into a plant cell;

culturing said cell to obtain descendant plant cells or plants comprising descendant plant cells, said descendant plant cells comprising said synthetic *B.t.* gene; and

establishing that said synthetic *B.t.* gene is expressed in said descendant plant cells.

2. The synthetic *Bacillus thuringiensis* gene of claim 1, comprising an A+T base content of less than about 60%.

3. The synthetic *Bacillus thuringiensis* gene of claim 1, wherein the process of producing said synthetic gene contains the additional step of determining the frequency of codon usage of said modified coding sequence.

4. A method of designing a synthetic *Bacillus thuringiensis* (*B.t.*) gene which is expressed in descendant plant cells, comprising the steps of:

selecting a *B.t.* pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native *B.t.* gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native *B.t.* coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native *B.t.* coding sequence;

obtaining a synthetic *B.t.* gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic *B.t.* gene into a plant cell;

culturing said cell to obtain descendant plant cells, said descendant plant cells comprising said synthetic *B.t.* gene; and

establishing that said synthetic *B.t.* gene is expressed in said descendant plant cells.

5. The method of claim 4, wherein said synthetic gene comprises an A+T base content of less than about 60%.

6. The method of claim 4, wherein said method further comprises the step of determining the frequency of codon usage of said modified coding sequence.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,015,891
DATED : January 18, 2000
INVENTOR(S) : Michael J. Adang, Elizabeth E. Murray

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 4,
Line 42: "FIGS. 1A and 1B" should read -- Figures 1A, 1B and 1C --;
Line 58: "liated" should read -- ligated --.

Column 5,
Line 67: "nay" should read -- may --.

Column 7,
Line 30: "call" should read -- cell --.

Column 11,
Line 15: "inonocots" should read -- monocots --;
Line 23: "nonocots" should read -- monocots --.

Column 15,
Line 27: "illustrations. of" should read -- illustrations of --.

Signed and Sealed this

Eleventh Day of September, 2001

Attest:

Nicholas P. Godici

Attesting Officer

NICHOLAS P. GODICI
Acting Director of the United States Patent and Trademark Office

APPENDIX B

US006013523A

United States Patent [19]

Adang et al.

[11] Patent Number: 6,013,523

[45] Date of Patent: *Jan. 11, 2000

[54] **TRANSGENIC PLANTS COMPRISING A SYNTHETIC INSECTICIDAL CRYSTAL PROTEIN GENE HAVING A MODIFIED FREQUENCY OF CODON USAGE**

[75] Inventors: **Michael J. Adang; Elizabeth E. Murray**, both of Madison, Wis.

[73] Assignee: **Mycogen Plant Science, Inc.**, San Diego, Calif.

[*] Notice: This patent is subject to a terminal disclaimer.

[21] Appl. No.: **08/704,966**

[22] Filed: **Aug. 29, 1996**

Related U.S. Application Data

[62] Division of application No. 08/369,839, Jan. 6, 1995, Pat. No. 5,567,862, which is a division of application No. 08/057,191, May 3, 1993, Pat. No. 5,380,831, which is a continuation of application No. 07/827,844, Jan. 28, 1992, abandoned, which is a continuation of application No. 07/242,482, Sep. 9, 1988, abandoned.

[51] **Int. Cl.**[7] **C12N 5/14**

[52] **U.S. Cl.** **435/419**; 536/23.71

[58] **Field of Search** 536/23.71; 435/172.3, 435/320.1, 69.1, 419, 468, 440; 800/205, 279, 302

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,356,270	10/1982	Itakura	435/317
4,888,282	12/1989	Beremand et al.	435/193
5,254,799	10/1993	De Greve	800/205
5,276,268	1/1994	Strauch et al.	800/205
5,380,831	1/1995	Adang et al.	536/23.71
5,567,862	10/1996	Adang et al.	800/205

FOREIGN PATENT DOCUMENTS

0305275	3/1989	European Pat. Off. .

OTHER PUBLICATIONS

Schnepf HE. "*Bacillus thruingiensis* toxins: Regulation, activities and structural diversity." Current Opinion Biotech. 6: 305–312, 1995.

Sekar, V. et al. (1987) "Molecular cloning and characterization of the insecticidal crystal protein gene of *Bacillus thuringiensis* var. *tenebrionis*" Pro. Natl. Acad. Sci. USA 84:7036–7040.

McPherson, S. A. et al. (1988) "Characterization of the Colopteran–specific protein gene of *Bacillus thuringiensis* var. *Tenebrionis*" Biotechnology 6(1):61–6.

Vaeck, M. et al. (1987) "Transgenic Plants protected from insect attach" Nature 328:33–37.

Murray, E. E. et al. (1989) "Codon usage in plant genes" Nucleic Acids Research 17(2):477–498.

Perlak, F. J. et al. (1991) "Modification of the coding sequence enhances plant expression of insect control protein genes" Proc. Natl. Acad. Sci. USA 88:3324–3328.

Hernastadt et al. (1986) Biotechnology 4:305–308.

Sekar et al. (1987) Proc. Natl. Acad. Sci. 84:7036–7040.

McPherson et al. (1988) Biotechnology 6:61–66.

Vaeck et al. (1987) Nature 328:33–37.

Barton et al. (1987) Plant Physiol. 85: 1103–1109.

Hofte et al. (1988) FEBS Lett. 226:364–370.

Bennetzen and Hall (1982) J. Biol. Chem. 257:3026–3031.

Hoekema et al. (1987) Mol. Cell. Biol. 7:2914–2924.

Boudraa (1987) Genet. Sel. Evol. 19:143–154.

Grantham et al. (1986) Oxford Surveys in Evol. Biol. 3:48–81.

Taylor et al. (1987) Mol. Gen. Genet. 210:572–577.

Joshi (1987) Nucl. Acids Res. 15:9627–9640.

Vankan and Fillipowicz (1988) EMBO J. 7:791–799.

Tuerk et al. (1988) Proc. Natl. Acad. Sci. 85:1364–1368.

Krieg et al. (1983) Z. Angew. Entomol. 96:500–508.

Bernahard (1986) FEMS Microbiol. Lett. 33:261–265.

Neill et al. (1987) Gene 55:303–317.

Rothstein et al. (1987) Gene 55: 353–356.

Coraggio et al. (1986) EMBO J. 5:459–465.

Fuzakawa et al. (1987) FEBS Lett. 224: 125–127.

Vies et al. (1986) EMBO J. 5:2439–2444.

Gatenby et al. (1987) Eur. J. Biochem. 168: 227–231.

Viotti et al. (1978) Biochim. Biophys. Acta 517: 125–132.

Hames and Higgins (eds) (1985) *Nucleic Acid Hybridisation*, IRL Press, Oxford, UKe.

Caruthers, M. (1983) In Methodology of DNA and RNA Sequencing, Chapter 1, Weissman (ed.), Praeger Publishers, New Yor.l.

Lo et al. (1984) Proc. Natl. Acad. Sci. 81: 2285–2289.

Grantham et al. (1985) Bull. inst. Pasteur 83: 95–148.

Matsuoka et al. (1987) J. Biochem. 102:673–676.

Mandecki et al. (1985) Proc. Natl. Acad. Sci. 82: 3543–3547.

Feretti et al. (1986) Proc. Natl. Sci. 83:599–603.

Paszkowski, J. et al. (1984) EMBO J. 3: 2717.

Fromm, M. et al. (1985) Proc. Natl. Acad. Sci. USA 82:5824.

Crossway, A. et al. (1986) Mol. Gen. Genet. 202:179.

Hooykaas–Van Slogteren, G. et al. (1984) Nature 311:763.

Dandekar, A. et al. (1987) Biotechnology 5:587.

An, G. et al. (1985) EMBO J. 4:277.

Herrera–Estrella, L. et al. (1983) Nature 303:209E.

Herrera–Estrella, L. et al. (1983) EMBO J. 2:987.

Herrera–Estrella, L. et al. (1985) in Plant Genetic Engineering, Cambridge University Press, New York, p. 63.

asadaban, M. J. and Cohen, S.N. (1980) J. Mol. Biol. 138:179–207.

Marsh, J. L. et al. (1984) Gene 32: 481–485.

Kronstad et al. (1983) J. Bacteriol. 154: 419–428.

Adang et al. (1987) in Biotechnology in Invertebrate Pathology and Cell Culture, K. Maramorosh (ed.) Academic Press, Inc. New York, pp. 85–99.

(List continued on next page.)

Primary Examiner—Lynette R. F. Smith
Assistant Examiner—Amy J. Nelson
Attorney, Agent, or Firm—Saliwanchik, Lloyd & Saliwanchik

[57] **ABSTRACT**

Synthetic *Baccilus thuringiensis* toxin genes designed to be expressed in plants at a level higher than naturally-occurring *Bt* genes are provided. These genes utilize codons preferred in highly expressed monocot or dicot proteins.

4 Claims, 5 Drawing Sheets

OTHER PUBLICATIONS

Schnepf and Whitteley (1985) J. Biol. Chem 260:6273–6280.
Pfitzinger et al. (1987) Nucl. Acids Res. 15: 1377–1386.
Sharp and Li (1986) Nucl. Acids Res. 14:7734–7749.
Matteucci et al. (1981) J. Am. Chem. Soc. 103:3185–3192.
Beaucage et al. (1981) tetrahedron Lett. 22: 1859–1862.
McBride et al. (1988) Biotechniques 6:362–367.
Maniatis (1982) Cloning Manual, Fritsch and Sambrook (eds.), Cold Spring Harbor Laboratory, Cold Spring Harebor, New York.

Sanger et al. (1977) proc. Natl. Acad. Sci. 74: 5463–5467.

Russell and Bennett (1982) Gene 20: 231–244.

Yanisch–Perron, C. et al. (1985) Gene 33: 103–119.

Barnes (1986) J. Cell Biochem. O (10 part C) 47.

Boswell et al. in Computational Molecular Biology, Sources and Methods for Sequence Analysis (Leak, ed.) Oxford University Press, Oxford, 1986, pp. 170–171.

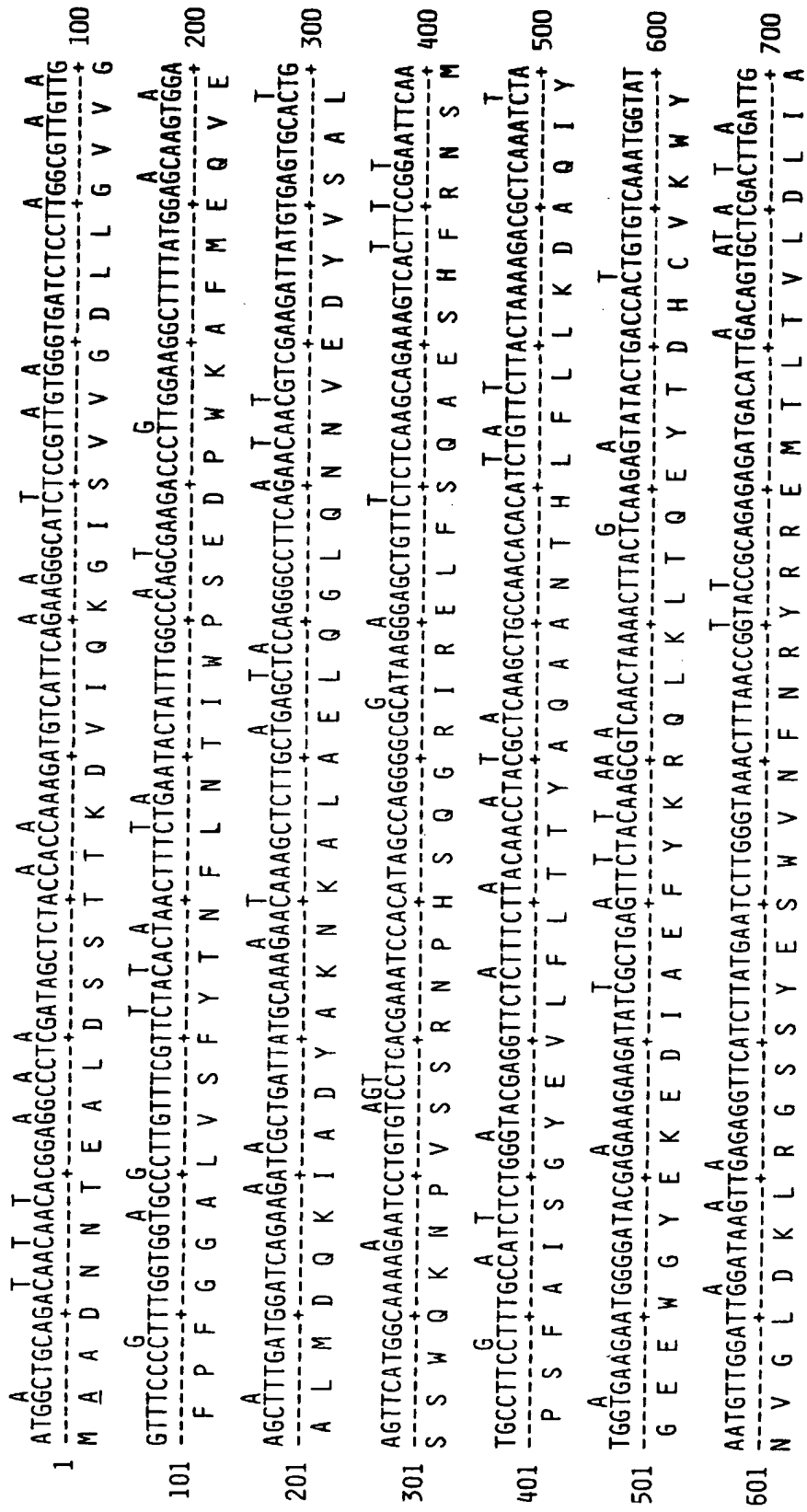


FIG.1A

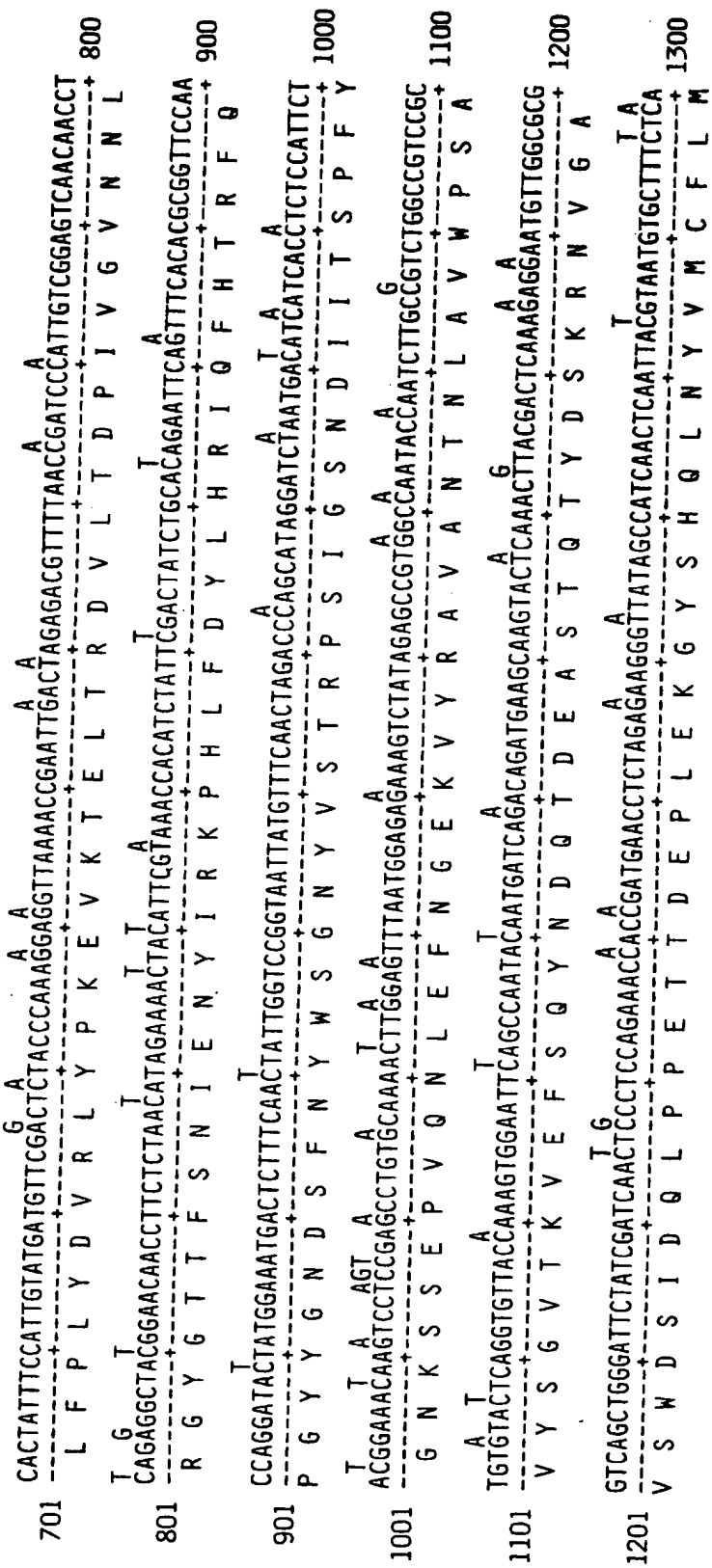


FIG. 1B

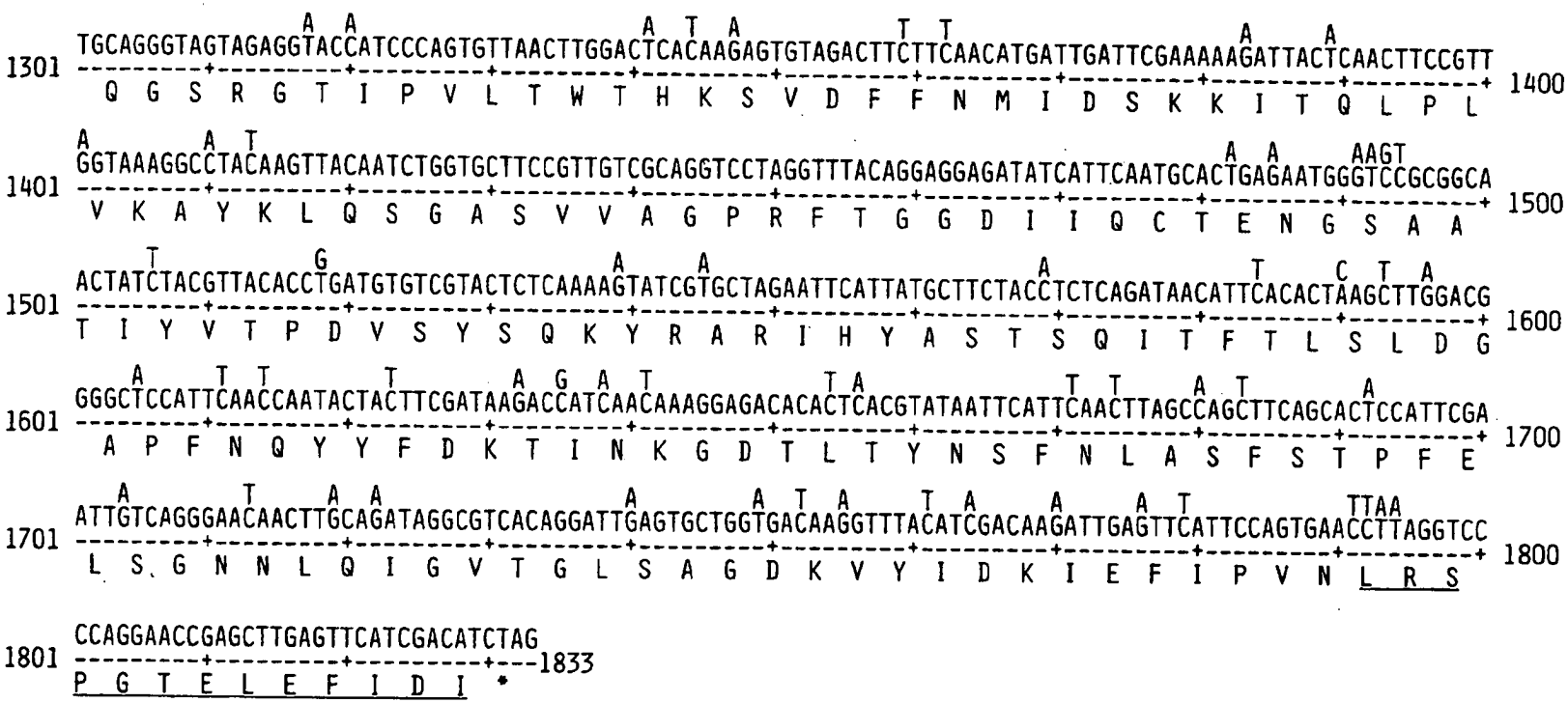

1301 TGCAGGGTAGTAGAGGTACCATCCCAGTGTTAACTTGGACTCACAAGAGTGTAGACTTCTTCAACATGATTGATTCGAAAAAGATTACTCAACTTCCGTT 1400
Q G S R G T I P V L T W T H K S V D F F N M I D S K K I T Q L P L
1401 GGTAAAGGCCTACAAGTTACAATCTGGTGCTTCCGTTGTCGCAGGTCCTAGGTTTACAGGAGGAGATATCATTCAATGCACTGAGAATGGGTCCGCGGCA 1500
V K A Y K L Q S G A S V V A G P R F T G G D I I Q C T E N G S A A
1501 ACTATCTACGTTACACCTGATGTGTCGTACTCTCAAAAGTATCGTGCTAGAATTCATTATGCTTCTACCTCTCAGATAACATTCACACTAAGCTTGGACG 1600
T I Y V T P D V S Y S Q K Y R A R I H Y A S T S Q I T F T L S L D G
1601 GGGCTCCATTCAACCAATACTACTTCGATAAGACCATCAACAAAGGAGACACACTCACGTATAATTCATTCAACTTAGCCAGCTTCAGCACTCCATTCGA 1700
A P F N Q Y Y F D K T I N K G D T L T Y N S F N L A S F S T P F E
1701 ATTGTCAGGGAACAACTTGCAGATAGGCGTCACAGGATTGAGTGCTGGTGACAAGGTTTACATCGACAAGATTGAGTTCATTCCAGTGAACCTTAGGTCC 1800
L S G N N L Q I G V T G L S A G D K V Y I D K I E F I P V N L R S
1801 CCAGGAACCGAGCTTGAGTTCATCGACATCTAG 1833
P G T E L E F I D I *

FIG. IC

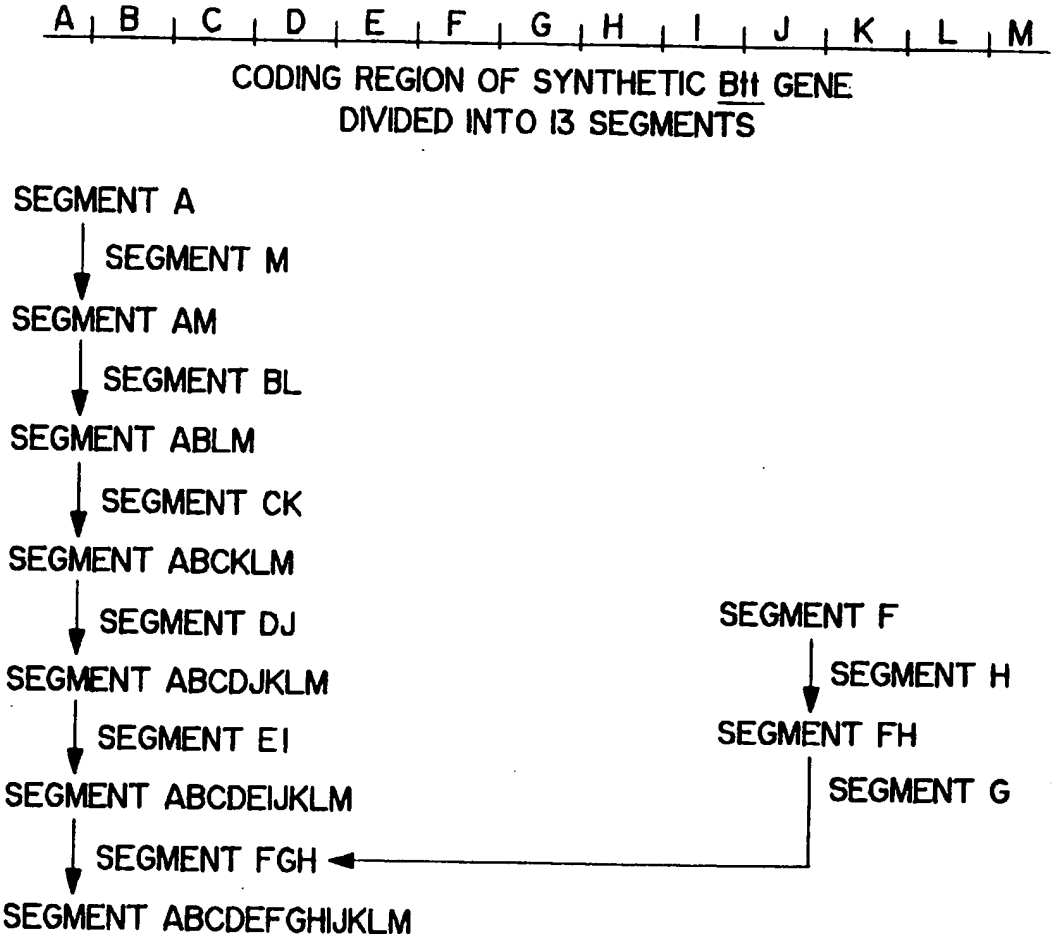

A B C D E F G H I J K L M
CODING REGION OF SYNTHETIC Bt1 GENE
DIVIDED INTO 13 SEGMENTS
SEGMENT A
SEGMENT M
SEGMENT AM
SEGMENT BL
SEGMENT ABLM
SEGMENT CK
SEGMENT ABCKLM
SEGMENT DJ
SEGMENT ABCDJKLM
SEGMENT EI
SEGMENT ABCDEIJKLM
SEGMENT FGH
SEGMENT ABCDEFGHIJKLM
SEGMENT F
SEGMENT H
SEGMENT FH
SEGMENT G

FIG. 2

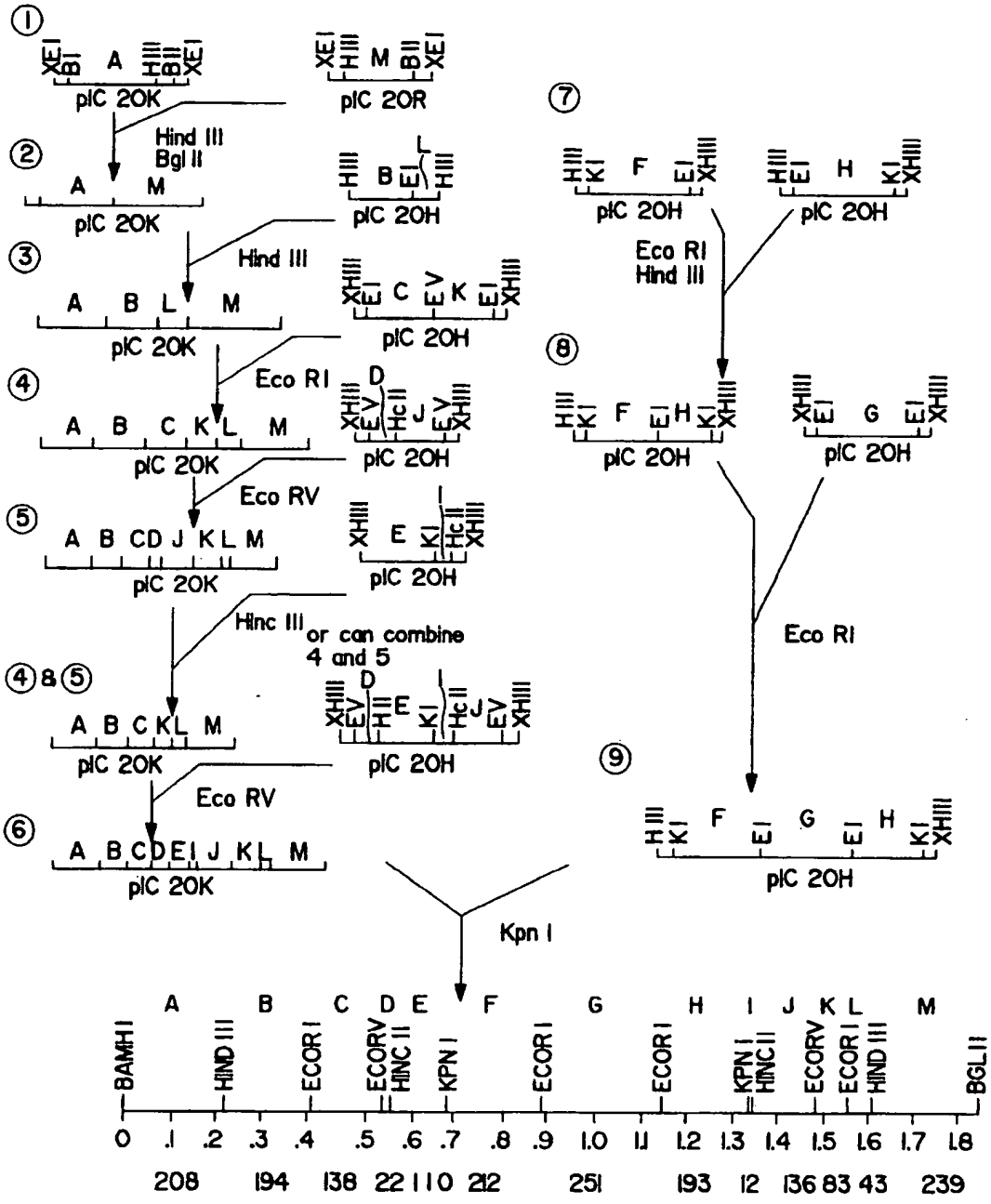

pIC 20K
pIC 20R
pIC 20H
Hind III
Bgl II
Eco RI
Eco RV
Hinc III
or can combine 4 and 5
Kpn I
BAMH I
HIND III
ECOR I
ECORV
HINC II
KPN I
BGL II
208 194 138 22 110 212 251 193 12 136 83 43 239

FIG. 3

TRANSGENIC PLANTS COMPRISING A SYNTHETIC INSECTICIDAL CRYSTAL PROTEIN GENE HAVING A MODIFIED FREQUENCY OF CODON USAGE

CROSS REFERENCES TO RELATED APPLICATIONS

This is a division of application Ser. No. 08/369,839, filed Jan. 6, 1995, now U.S. Pat. No. 5,567,862, which is a divisional of application Ser. No. 08/057,191, filed May 3, 1993 now U.S. Pat. No. 5,380,831; which is a continuation of application Ser. No. 07/827,844, filed Jan. 28, 1992, now abandoned; which is a continuation of application Ser. No. 07/242,482, filed Sep. 9, 1988, now abandoned, all of which are incorporated herein by reference.

FIELD OF THE INVENTION

This invention relates to the field of bacterial molecular biology and, in particular, to genetic engineering by recombinant technology for the purpose of protecting plants from insect pests. Disclosed herein are the chemical synthesis of a modified crystal protein gene from *Bacillus thuringiensis* var. *tenebrionis* (*Btt*), and the selective expression of this synthetic insecticidal gene. Also disclosed is the transfer of the cloned synthetic gene into a host microoorganism, rendering the organism capable of producing, at improved levels of expression, a protein having toxicity to insects. This invention facilitates the genetic engineering of bacteria and plants to attain desired expression levels of novel toxins having agronomic value.

BACKGROUND OF THE INVENTION

B. thuringiensis (*Bt*) is unique in its ability to produce, during the process of sporulation, proteinaceous, crystalline inclusions which are found to be highly toxic to several insect pests of agricultural importance. The crystal proteins of different *Bt* strains have a rather narrow host range and hence are used commercially as very selective biological insecticides. Numerous strains of *Bt* are toxic to lepidopteran and dipteran insects. Recently two subspecies (or varieties) of *Bt* have been reported to be pathogenic to coleopteran insects: var. *tenebrionis* (Krieg et al. (1983) Z. Angew. Entomol. 96:500–508) and var. *san diego* (Herrnstadt et al. (1986) Biotechnol. 4:305–308). Both strains produce flat, rectangular crystal inclusions and have a major crystal component of 64–68 kDa (Herrnstadt et al. supra; Bernhard (1986) FEMS Microbiol. Lett. 33:261–265).

Toxin genes from several subspecies of *Bt* have been cloned and the recombinant clones were found to be toxic to lepidopteran and dipteran insect larvae. The two coleopteran-active toxin genes have also been isolated and expressed. Herrnstadt et al. supra cloned a *5.8* kb BamHI fragment of *Bt* var. *san diego* DNA. The protein expressed in *E. coli* was toxic to *P. luteola* (Elm leaf beetle) and had a molecular weight of approximately 83 kDa. This 83 kDa toxin product from the var. *san diego* gene was larger than the 64 kDa crystal protein isolated from *Bt* var. *san diego* cells, suggesting that the *Bt* var. *san diego* crystal protein may be synthesized as a larger precursor molecule that is processed by *Bt* var. *san diego* but not by *E. coli* prior to being formed into a crystal.

Sekar et al. (*1987*) Proc. Nat. Acad. Sci. USA 84:7036–7040; U.S. patent application Ser. No. 108,285, filed Oct. 13, 1987 isolated the crystal protein gene from *Btt* and determined the nucleotide sequence. This crystal protein gene was contained on a 5.9 kb BamHI fragment (pNSBF544). A subclone containing the 3 kb HindIII fragment from pNSBF544 was constructed. This HindIII fragment contains an open reading frame (ORF) that encodes a 644-amino acid polypeptide of approximately 73 kDa. Extracts of both subclones exhibited toxicity to larvae of Colorado potato beetle (*Leptinotarsa decemlineata*, a coleopteran insect). *73*- and 65-kDa peptides that cross-reacted with an antiserum against the crystal protein of var. *tenebrionis* were produced on expression in *E. coli*. Sporulating var. *tenebrionis* cells contain an immunoreactive 73-kDa peptide that corresponds to the expected product from the ORF of pNSBP544. However, isolated crystals primarily contain a 65-kDa component. When the crystal protein gene was shortened at the N-terminal region, the dominant protein product obtained was the 65-kDa peptide. A deletion derivative, p544Pst-Met5, was enzymatically derived from the 5.9 kb BamHI fragment upon removal of forty-six amino acid residues from the N-terminus. Expression of the N-terminal deletion derivative, p544Pst-Met5, resulted in the production of, almost exclusively, the 65 kDa protein. Recently, McPherson et al. (1988) Biotechnology 6:61–66 demonstrated that the *Btt* gene contains two functional translational initiation codons in the same reading frame leading to the production of both the full-length protein and an N-terminal truncated form.

Chimeric toxin genes from several strains of *Bt* have been expressed in plants. Four modified *Bt2* genes from var. *berliner* 1715, under the control of the 2' promoter of the Agrobacterium TR-DNA, were transferred into tobacco plants (Vaeck et al. (1987) Nature 328:33–37). Insecticidal levels of toxin were produced when truncated genes were expressed in transgenic plants. However, the steady state mRNA levels in the transgenic plants were so low that they could not be reliably detected in Northern blot analysis and hence were quantified using ribonuclease protection experiments. *Bt* mRNA levels in plants producing the highest level of protein corresponded to ≈0.0001% of the poly(A)$^+$ mRNA.

In the report by Vaeck et al. (1987) supra, expression of chimeric genes containing the entire coding sequence of *Bt2* were compared to those containing truncated *Bt2* genes. Additionally, some T-DNA constructs included a chimeric NPTII gene as a marker selectable in plants, whereas other constructs carried translational fusions between fragments of *Bt2* and the NPTII gene. Insecticidal levels of toxin were produced when truncated *Bt2* genes or fusion constructs were expressed in transgenic plants. Greenhouse grown plants produced ≈0.02% of the total soluble protein as the toxin, or 3 μg of toxin per g. fresh leaf tissue and, even at five-fold lower levels, showed 100% mortality in six-day feeding assays. However, no significant insecticidal activity could be obtained using the intact *Bt2* coding sequence, despite the fact that the same promoter was used to direct its expression. Intact *Bt2* protein and RNA yields in the transgenic plant leaves were 10–50 times lower than those for the truncated *Bt2* polypeptides or fusion proteins.

Barton et al. (1987) Plant Physiol. 85:1103–1109 showed expression of a *Bt* protein in a system containing a 35S promoter, a viral (TMV) leader sequence, the *Bt* HD-1 4.5 kb gene (encoding a 645 amino acid protein followed by two proline residues) and a nopaline synthase (nos) poly(A)+ sequence. Under these conditions expression was observed for *Bt* mRNA at levels up to 47 pg/20 μg RNA and 12 ng/mg plant protein. This amount of *Bt* protein in plant tissue produced 100% mortality in two days. This level of expres-

sion still represents a low level of mRNA (2.5×10^{-4}%) and protein (1.2×10^{-3}%).

Various hybrid proteins consisting of N-terminal fragments of increasing length of the *Bt*2 protein fused to NPTII were produced in *E. coli* by Hofte et al. (1988) FEBS Lett. 226:364–370. Fusion proteins containing the first 607 amino acids of *Bt*2 exhibited insect toxicity; fusion proteins not containing this minimum N-terminal fragment were nontoxic. Appearance of NPTII activity was not dependent upon the presence of insecticidal activity; however, the conformation of the *Bt*2 polypeptide appeared to exert an important influence on the enzymatic activity of the fused NPTII protein. This study did suggest that the global 3-D structure of the *Bt*2 polypeptide is disturbed in truncated polypeptides.

A number of researchers have attempted to express plant genes in yeast (Neill et al. (1987) Gene 55:303–317; Rothstein et al. (1987) Gene 55:353–356; Coraggio et al. (1986) EMBO J. 5:459–465) and *E. coli* (Fuzakawa et al. (1987) FEBS Lett. 224:125–127; Vies et al. (1986) EMBO J. 5:2439–2444; Gatenby et al. (1987) Eur. J. Biochem. 168:227–231). In the case of wheat α-gliadin (Neill et al. (1987) supra); α-amylase (Rothstein et al. (1987) supra) genes, and maize zein genes (Coraggio et al. (1986) supra) in yeast, low levels of expression have been reported. Neill et al. have suggested that the low levels of expression of α-gliadin in yeast may be due in part to codon usage bias, since α-gliadin codons for Phe, Leu, Ser, Gly, Tyr and especially Glu do not correlate well with the abundant yeast isoacceptor tRNAs. In *E. coli* however, soybean glycinin A2 (Fuzakawa et al. (1987) supra) and wheat RuBPC SSU (Vies et al. (1986) supra; Gatenby et al. (1987) supra) are expressed adequately.

Not much is known about the makeup of tRNA populations in plants. Viotti et al. (1978) Biochim. Biophys. Acta 517:125–132 report that maize endosperm actively synthesizing zein, a storage protein rich in glutamine, leucine, and alanine, is characterized by higher levels of accepting activity for these three amino acids than are maize embryo tRNAs. This may indicate that the tRNA population of specific plant tissues may be adapted for optimum translation of highly expressed proteins such as zein. To our knowledge, no one has experimentally altered codon bias in highly expressed plant genes to determine possible effects of the protein translation in plants to check the effects on the level of expression.

SUMMARY OF THE INVENTION

It is the overall object of the present invention to provide a means for plant protection against insect damage. The invention disclosed herein comprises a chemically synthesized gene encoding an insecticidal protein which is functionally equivalent to a native insecticidal protein of *Bt*. This synthetic gene is designed to be expressed in plants at a level higher than a native *Bt* gene. It is preferred that the synthetic gene be designed to be highly expressed in plants as defined herein. Preferably, the synthetic gene is at least approximately 85% homologous to an insecticidal protein gene of *Bt*.

It is a particular object of this invention to provide a synthetic structural gene coding for an insecticidal protein from *Btt* having, for example, the nucleotide sequences presented in FIG. 1 and spanning nucleotides 1 through 1793 or spanning nucleotide 1 through 1833 with functional equivalence.

In designing synthetic *Btt* genes of this invention for enhanced expression in plants, the DNA sequence of the native *Btt* structural gene is modified in order to contain codons preferred by highly expressed plant genes, to attain an A+T content in nucleotide base composition substantially that found in plants, and also preferably to form a plant initiation sequence, and to eliminate sequences that cause destabilization, inappropriate polyadenylation, degradation and termination of RNA and to avoid sequences that constitute secondary structure hairpins and RNA splice sites. In the synthetic genes, codons used to specify a given amino acid are selected with regard to the distribution frequency of codon usage employed in highly expressed plant genes to specify that amino acid. As is appreciated by those skilled in the art, the distribution frequency of codon usage utilized in the synthetic gene is a determinant of the level of expression. Hence, the synthetic gene is designed such that its distribution frequency of codon usage deviates, preferably, no more than 25% from that of highly expressed plant genes and, more preferably, no more than about 10%. In addition, consideration is given to the percentage G+C content of the degenerate third base (monocotyledons appear to favor G+C in this position, whereas dicotyledons do not). It is also recognized that the XCG nucleotide is the least preferred codon in dicots whereas the XTA codon is avoided in both monocots and dicots. The synthetic genes of this invention also preferably have CG and TA doublet avoidance indices as defined in the Detailed Description closely approximating those of the chosen host plant. More preferably these indices deviate from that of the host by no more than about 10–15%.

Assembly of the *Bt* gene of this invention is performed using standard technology known to the art. The *Btt* structural gene designed for enhanced expression in plants of the specific embodiment is enzymatically assembled within a DNA vector from chemically synthesized oligonucleotide duplex segments. The synthetic *Bt* gene is then introduced into a plant host cell and expressed by means known to the art. The insecticidal protein produced upon expression of the synthetic *Bt* gene in plants is functionally equivalent to a native *Bt* crystal protein in having toxicity to the same insects.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 presents the nucleotide sequence for the synthetic *Btt* gene. Where different, the native sequence as found in p544Pst-Met5 is shown above. Changes in amino acids (underlined) occur in the synthetic sequence with alanine replacing threonine at residue 2 and leucine replacing the stop at residue 596 followed by the addition of 13-amino acids at the C-terminus.

FIG. 2 represents a simplified scheme used in the construction of the synthetic *Btt* gene. Segments A through M represent oligonucleotide pieces annealed and ligated together to form DNA duplexes having unique splice sites to allow specific enzymatic assembly of the DNA segments to give the desired gene.

FIG. 3 is a schematic diagram showing the assembly of oligonucleotide segments in the construction of a synthetic *Btt* gene. Each segment (A through M) is built from oligonucleotides of different sizes, annealed and ligated to form the desired DNA segment.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO. 1 is the native DNA gene sequence corresponding to FIG. 1.

SEQ ID NO. 2 is the corresponding amino acid sequence encoded by SEQ ID NO. 1.

SEQ ID NO. 3 is the synthetic DNA gene sequence corresponding to FIG. 1.

SEQ ID NO. 4 is the corresponding amino acid sequence encoded by SEQ ID NO. 3.

SEQ ID NO. 5 is the sequence of a synthetic DNA linker described in Example 1(i).

SEQ ID NO. 6 is the sequence of the 5' plant consensus splice site found in Example 1(iii)(d).

SEQ ID NO. 7 is the sequence of the 3' plant consensus splice site found in Example 1 (iii)(d).

SEQ ID NO. 8 is the nucleotide sequence of Segment A found in Table 4.

SEQ ID NO. 9 is the nucleotide sequence of Segment M found in Table 5.

DETAILED DESCRIPTION OF THE INVENTION

The following definitions are provided in order to provide clarity as to the intent or scope of their usage in the Specification and claims.

Expression refers to the transcription and translation of a structural gene to yield the encoded protein. The synthetic *Bt* genes of the present invention are designed to be expressed at a higher level in plants than the corresponding native *Bt* genes. As will be appreciated by those skilled in the art, structural gene expression levels are affected by the regulatory DNA sequences (promoter, polyadenylation sites, enhancers, etc.) employed and by the host cell in which the structural gene is expressed. Comparisons of synthetic *Bt* gene expression and native *Bt* gene expression must be made employing analogous regulatory sequences and in the same host cell. It will also be apparent that analogous means of assessing gene expression must be employed in such comparisons.

Promoter refers to the nucleotide sequences at the 5' end of a structural gene which direct the initiation of transcription. Promoter sequences are necessary, but not always sufficient, to drive the expression of a downstream gene. In prokaryotes, the promoter drives transcription by providing binding sites to RNA polymerases and other initiation and activation factors. Usually promoters drive transcription preferentially in the downstream direction, although promotional activity can be demonstrated (at a reduced level of expression) when the gene is placed upstream of the promoter. The level of transcription is regulated by promoter sequences. Thus, in the construction of heterologous promoter/structural gene combinations, the structural gene is placed under the regulatory control of a promoter such that the expression of the gene is controlled by promoter sequences. The promoter is positioned preferentially upstream to the structural gene and at a distance from the transcription start site that approximates the distance between the promoter and the gene it controls in its natural setting. As is known in the art, some variation in this distance can be tolerated without loss of promoter function.

A gene refers to the entire DNA portion involved in the synthesis of a protein. A gene embodies the structural or coding portion which begins at the 5' end from the translational start codon (usually ATG) and extends to the stop (TAG, TGA or TAA) codon at the 3' end. It also contains a promoter region, usually located 5' or upstream to the structural gene, which initiates and regulates the expression of a structural gene. Also included in a gene are the 3' end and poly(A)+ addition sequences.

Structural gene is that portion of a gene comprising a DNA segment encoding a protein, polypeptide or a portion thereof, and excluding the 5' sequence which drives the initiation of transcription. The structural gene may be one which is normally found in the cell or one which is not normally found in the cellular location wherein it is introduced, in which case it is termed a heterologous gene. A heterologous gene may be derived in whole or in part from any source know to the art, including a bacterial genome or episome, eukaryotic, nuclear or plasmid DNA, cDNA, viral DNA or chemically synthesized DNA. A structural gene nay contain one or more modifications in either the coding or the untranslated regions which could affect the biological activity or the chemical structure of the expression product, the rate of expression or the manner of expression control. Such modifications include, but are not limited to, mutations, insertions, deletions and substitutions of one or more nucleotides. The structural gene may constitute an uninterrupted coding sequence or it may include one or more introns, bounded by the appropriate splice junctions. The structural gene may be a composite of segments derived from a plurality of sources, naturally occurring or synthetic. The structural gene may also encode a fusion protein.

Synthetic gene refers to a DNA sequence of a structural gene that is chemically synthesized in its entirety or for the greater part of the coding region. As exemplified herein, oligonucleotide building blocks are synthesized using procedures known to those skilled in the art and are ligated and annealed to form gene segments which are then enzymatically assembled to construct the entire gene. As is recognized by those skilled in the art, functionally and structurally equivalent genes to the synthetic genes described herein may be prepared by site-specific mutagenesis or other related methods used in the art.

Transforming refers to stably introducing a DNA segment carrying a functional gene into an organism that did not previously contain that gene.

Plant tissue includes differentiated and undifferentiated tissues of plants, including but not limited to, roots, shoots, leaves, pollen, seeds, tumor tissue and various forms of cells in culture, such as single cells, protoplasts, embryos and callus tissue. The plant tissue may be in planta or in organ, tissue or cell culture.

Plant cell as used herein includes plant cells in planta and plant cells and protoplasts in culture.

Homology refers to identity or near identity of nucleotide or amino acid sequences. As is understood in the art, nucleotide mismatches can occur at the third or wobble base in the codon without causing amino acid substitutions in the final polypeptide sequence. Also,. minor nucleotide modifications (e.g., substitutions, insertions or deletions) in certain regions of the gene sequence can be tolerated and considered insignificant whenever such modifications result in changes in amino acid sequence that do not alter functionality of the final product. It has been shown that chemically synthesized copies of whole, or parts of, gene sequences can replace the corresponding regions in the natural gene without loss of gene function. Homologs of specific DNA sequences may be identified by those skilled in the art using the test of cross-hybridization of nucleic acids under conditions of stringency as is well understood in the art (as described in Hames and Higgens (eds.) (1985) *Nucleic Acid Hybridization*, IRL Press, Oxford, UK). Extent of homology is often measured in terms of percentage of identity between the sequences compared.

Functionally equivalent refers to identity or near identity of function. A synthetic gene product which is toxic to at least one of the same insect species as a natural *Bt* protein is considered functionally equivalent thereto. As exemplified

herein, both natural and synthetic *Btt* genes encode 65 kDa insecticidal proteins having essentially identical amino acid sequences and having toxicity to coleopteran insects. The synthetic *Bt* genes of the present invention are not considered to be functionally equivalent to native *Bt* genes, since they are expressible at a higher level in plants than native *Bt* genes.

Frequency of preferred codon usage refers to the preference exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. To determine the frequency of usage of a particular codon in a gene, the number of occurrences of that codon in the gene is divided by the total number of occurrences of all codons specifying the same amino acid in the gene. Table 1, for example, gives the frequency of codon usage for *Bt* genes, which was obtained by analysis of four *Bt* genes whose sequences are publicly available. Similarly, the frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell. It is preferable that this analysis be limited to genes that are highly expressed by the host cell. Table 1, for example, gives the frequency of codon usage by highly expressed genes exhibited by dicotyledonous plants, and monocotyledonous plants. The dicot codon usage was calculated using 154 highly expressed coding sequences obtained from Genbank which are listed in Table 1. Monocot codon usage was calculated using 53 monocot nuclear gene coding sequences obtained from Genbank and listed in Table 1, located in Example 1.

When synthesizing a gene for improved expression in a host cell it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

The percent deviation of the frequency of preferred codon usage for a synthetic gene from that employed by a host cell is calculated first by determining the percent deviation of the frequency of usage of a single codon from that of the host cell followed by obtaining the average deviation over all codons. As defined herein this calculation includes unique codons (i.e., ATG and TGG). The frequency of preferred codon usage of the synthetic *Btt* gene, whose sequence is given in FIG. 1, is given in Table 1. The frequency of preferred usage of the codon 'GTA' for valine in the synthetic gene (0.10) deviates from that preferred by dicots (0.12) by 0.02/0.12=0.167 or 16.7%. The average deviation over all amino acid codons of the *Btt* synthetic gene codon usage from that of dicot plants is 7.8%. In general terms the overall average deviation of the codon usage of a synthetic gene from that of a host cell is calculated using the equation

$$A = \sum_{n=1}^{Z} \frac{\frac{X_n - Y_n}{X_n} \times 100}{Z}$$

where X_n=frequency of usage for codon n in the host cell; Y_n=frequency of usage for codon n in the synthetic gene. Where n represents an individual codon that specifies an amino acid, the total number of codons is Z, which in the preferred embodiment is 61. The overall deviation of the frequency of codon usage, A, for all amino acids should preferably be less than about 25%, and more preferably less than about 10%.

Derived from is used to mean taken, obtained, received, traced, replicated or descended from a source (chemical and/or biological). A derivative may be produced by chemical or biological manipulation (including but not limited to substitution, addition, insertion, deletion, extraction, isolation, mutation and replication) of the original source.

Chemically synthesized, as related to a sequence of DNA, means that the component nucleotides were assembled in vitro. Manual chemical synthesis of DNA may be accomplished using well established procedures (Caruthers, M. (1983) in *Methodology of DNA and RNA Sequencing*, Weissman (ed.), Praeger Publishers, New York, Chapter 11), or automated chemical synthesis can be performed using one of a number of commercially available machines.

The term, designed to be highly expressed as used herein refers to a level of expression of a designed gene wherein the amount of its specific mRNA transcripts produced is sufficient to be quantified in Northern blots and, thus, represents a level of specific mRNA expressed corresponding to greater than or equal to approximately 0.001% of the poly(A)+ mRNA. To date, natural *Bt* genes are transcribed at a level wherein the amount of specific mRNA produced is insufficient to be estimated using the Northern blot technique. However, in the present invention, transcription of a synthetic *Bt* gene designed to be highly expressed not only allows quantification of the specific mRNA transcripts produced but also results in enhanced expression of the translation product which is measured in insecticidal bioassays.

Crystal protein or insecticidal crystal protein or crystal toxin refers to the major protein component of the parasporal crystals formed in strains of *Bt*. This protein component exhibits selective pathogenicity to different species of insects. The molecular size of the major protein isolated from parasporal crystals varies depending on the strain of *Bt* from which it is derived. Crystal proteins having molecular weights of approximately 132, 65, and 28 kDa have been reported. It has been shown that the approximately 132 kDa protein is a protoxin that is cleaved to form an approximately 65 kDa toxin.

The crystal protein gene refers to the DNA sequence encoding the insecticidal crystal protein in either full length protoxin or toxin form, depending on the strain of *Bt* from which the gene is derived.

The authors of this invention observed that expression in plants of *Bt* crystal protein mRNA occurs at levels that are not routinely detectable in Northern blots and that low levels of *Bt* crystal protein expression correspond to this low level of mRNA expression. It is preferred for exploitation of these genes as potential biocontrol methods that the level of expression of *Bt* genes in plant cells be improved and that the stability of *Bt* mRNA in plants be optimized. This will allow greater levels of *Bt* mRNA to accumulate and will result in an increase in the amount of insecticidal protein in plant tissues. This is essential for the control of insects that are relatively resistant to *Bt* protein.

Thus, this invention is based on the recognition that expression levels of desired, recombinant insecticidal protein in transgenic plants can be improved via increased expression of stabilized mRNA transcripts; and that, conversely, detection of these stabilized RNA transcripts may be utilized to measure expression of translational product (protein). This invention provides a means of resolving the problem of low expression of insecticidal protein RNA in plants and, therefore, of low protein expression through the use of an improved, synthetic gene specifying an insecticidal crystal protein from *Bt*.

Attempts to improve the levels of expression of *Bt* genes in plants have centered on comparative studies evaluating parameters such as gene type, gene length, choice of promoters, addition of plant viral untranslated RNA leader,

addition of intron sequence and modification of nucleotides surrounding the initiation ATG codon. To date, changes in these parameters have not led to significant enhancement of *Bt* protein expression in plants. Applicants find that, surprisingly, to express *Bt* proteins at the desired level in plants, modifications in the coding region of the gene were effective. Structural-function relationships can be studied using site-specific mutagenesis by replacement of restriction fragments with synthetic DNA duplexes containing the desired nucleotide changes (Lo et al. (1984) Proc. Natl. Acad. Sci. 81:2285–2289). However, recent advances in recombinant DNA technology now make it feasible to chemically synthesize an entire gene designed specifically for a desired function. Thus, the *Btt* coding region was chemically synthesized, modified in such a way as to improve its expression in plants. Also, gene synthesis provides the opportunity to design the gene so as to facilitate its subsequent mutagenesis by incorporating a number of appropriately positioned restriction endonuclease sites into the gene.

The present invention provides a synthetic *Bt* gene for a crystal protein toxic to an insect. As exemplified herein, this protein is toxic to coleopteran insects. To the end of improving expression of this insecticidal protein in plants, this invention provides a DNA segment homologous to a *Btt* structural gene and, as exemplified herein, having approximately 85% homology to the *Btt* structural gene in p544Pst-Met5. In this embodiment the structural gene encoding a *Btt* insecticidal protein is obtained through chemical synthesis of the coding region. A chemically synthesized gene is used in this embodiment because it best allows for easy and efficacious accommodation of modifications in nucleotide sequences required to achieve improved levels of cross-expression.

Today, in general, chemical synthesis is a preferred method to obtain a desired modified gene. However, to date, no plant protein gene has been chemically synthesized nor has any synthetic gene for a bacterial protein been expressed in plants. In this invention, the approach adopted for synthesizing the gene consists of designing an improved nucleotide sequence for the coding region and assembling the gene from chemically synthesized oligonucleotide segments. In designing the gene, the coding region of the naturally-occurring gene, preferably from the *Btt* subclone, p544Pst-Met5, encoding a 65 kDa. polypeptide having coleoperan toxicity, is scanned for possible modifications which would result in improved expression of the synthetic gene in plants. For example, to optimize the efficiency of translation, codons preferred in highly expressed proteins of the host cell are utilized.

Bias in codon choice within genes in a single species appears related to the level of expression of the protein encoded by that gene. Codon bias is most extreme in highly expressed proteins of *E. coli* and yeast. In these organisms, a strong positive correlation has been reported between the abundance of an isoaccepting tRNA species and the favored synonymous codon. In one group of highly expressed proteins in yeast, over 96% of the amino acids are encoded by only. 25 of the 61 available codons (Bennetzen and Hall (1982) J. Biol. Chem. 257:3026–3031). These 25 codons are preferred in all sequenced yeast genes, but the degree of preference varies with the level of expression of the genes. Recently, Hoekema and colleagues (1987) Mol. Cell. Biol. 7:2914–2924 reported that replacement of these 25 preferred codons by minor codons in the 5' end of the highly expressed yeast gene PGK1 results in a decreased level of both protein and mRNA. They concluded that biased codon choice in highly expressed genes enhances translation and is required for maintaining mRNA stability in yeast. Without doubt, the degree of codon bias is an important factor to consider when engineering high expression of heterologous genes in yeast and other systems.

Experimental evidence obtained from point mutations and deletion analysis has indicated that in eukaryotic genes specific sequences are associated with post-transcriptional processing, RNA destabilization, translational termination, intron splicing and the like. These are preferably employed in the synthetic genes of this invention. In designing a bacterial gene for expression in plants, sequences which interfere with the efficacy of gene expression are eliminated.

In designing a synthetic gene, modifications in nucleotide sequence of the coding region are made to modify the A+T content in DNA base composition of the synthetic gene to reflect that normally found in genes for highly expressed proteins native to the host cell. Preferably the A+T content of the synthetic gene is substantially equal to that of said genes for highly expressed proteins. In genes encoding highly expressed plant proteins, the A+T content is approximately 55%. It is preferred that the synthetic gene have an A+T content near this value, and not sufficiently high as to cause destabilization of RNA and, therefore, lower the protein expression levels. More preferably, the A+T content is no more than about 60% and most preferably is about 55%. Also, for ultimate expression in plants, the synthetic gene nucleotide sequence is preferably modified to form a plant initiation sequence at the 5' end of the coding region. In addition, particular attention is preferably given to assure that unique restriction sites are placed in strategic positions to allow efficient assembly of oligonucleotide segments during construction of the synthetic gene and to facilitate subsequent nucleotide modification. As a result of these modifications in coding region of the native *Bt* gene, the preferred synthetic gene is expressed in plants at an enhanced level when compared to that observed with natural *Bt* structural genes.

In specific embodiments, the synthetic *Bt* gene of this invention encodes a *Btt* protein toxic to coleopteran insects. Preferably, the toxic polypeptide is about 598 amino acids in length, is at least 75% homologous to a *Btt* polypeptide, and, as exemplified herein, is essentially identical to the protein encoded by p544Pst-Met5, except for replacement of threonine by alanine at residue 2. This amino acid substitution results as a consequence of the necessity to introduce a guanine base at position +4 in the coding sequence.

In designing the synthetic gene of this invention, the coding region from the *Btt* subclone, p544Pst-Met5, encoding a 65 kDa polypeptide having coleopteran toxicity, is scanned for possible modifications which would result in improved expression of the synthetic gene in plants. For example, in preferred embodiments, the synthetic insecticidal protein is strongly expressed in dicot plants, e.g., tobacco, tomato, cotton, etc., and hence, a synthetic gene under these conditions is designed to incorporate to advantage codons used preferentially by highly expressed dicot proteins. In embodiments where enhanced expression of insecticidal protein is desired in a monocot, codons preferred by highly expressed monocot proteins (given in Table 1) are employed in designing the synthetic gene.

In general, genes within a taxonomic group exhibit similarities in codon choice, regardless of the function of these genes. Thus an estimate of the overall use of the genetic code by a taxonomic group can be obtained by summing codon frequencies of all its sequenced genes. This species-specific codon choice is reported in this invention from analysis of

208 plant genes. Both monocot and dicot plants are analyzed individually to determine whether these broader taxonomic groups are characterized by different patterns of synonymous codon preference. The 208 plant genes included in the codon analysis code for proteins having a wide range of functions and they represent 6 monocot and 36 dicot species. These proteins are present in different plant tissues at varying levels of expression.

In this invention it is shown that the relative use of synonymous codons differs between the monocots and the dicots. In general, the most important factor in discriminating between monocot and dicot patterns of codon usage is the percentage G+C content of the degenerate third base. In monocots, 16 of 18 amino acids favor G+C in this position, while dicots only favor G+C in 7 of 18 amino acids.

The G ending codons for Thr, Pro, Ala and Ser are avoided in both monocots and dicots because they contain C in codon position II. The CG dinucleotide is strongly avoided in plants (Boudraa (1987) Genet. Sel. Evol. 19:143–154) and other eukaryotes (Grantham et al. (1985) Bull. Inst. Pasteur 83:95–148), possibly due to regulation involving methylation. In dicots, XCG is always the least favored codon, while in monocots this is not the case. The doublet TA is also avoided in codon positions II and III in most eukaryotes, and this is true of both monocots and dicots.

Grantham and colleagues (1986) Oxford Surveys in Evol. Biol. 3:48–81 have developed two codon choice indices to quantify CG and TA doublet avoidance in codon positions II and III. XCG/XCC is the ratio of codons having C as base II of G-ending to C-ending triplets, while XTA/XTT is the ratio of A-ending to T-ending triplets with T as the second base. These indices have been calculated for the plant data in this paper (Table 2) and support the conclusion that monocot and dicot species differ in their use of these dinucleotides.

TABLE 2

Avoidance of CG and TA doublets in codons position II-III.
XCG/XCC and XTA/XAA values are multiplied by 100.

Group	Plants	Di-cots	Mono-cots	Maize	Soy-bean	RuBPC SSU	CAB
XCG/XCC	40	30	61	67	37	18	22
XTA/XTT	37	35	47	43	41	9	13

RuBPC SSU = ribulose 1,5 bisphosphate small subunit
CAB = chlorophyll a/b binding protein

Additionally, for two species, soybean and maize, species-specific codon usage profiles were calculated (not shown). The maize codon usage pattern resembles that of monocots in general, since these sequences represent over half of the monocot sequences available. The codon profile of the maize subsample is even more strikingly biased in its preference for G+C in codon position III. On the other hand, the soybean codon usage pattern is almost identical to the general dicot pattern, even though it represents a much smaller portion of the entire dicot sample.

In order to determine whether the coding strategy of highly expressed genes such as the ribulose 1,5 bisphosphate small subunit (RuBPC SSU) and chlorophyll a/b binding protein (CAB) is more biased than that of plant genes in general, codon usage profiles for subsets of these genes (19 and 17 sequences, respectively) were calculated (not shown). The RuBPC SSU and CAB pooled samples are characterized by stronger avoidance of the codons XCG and XTA than in the larger monocot and dicot samples (Table 2). Although most of the genes in these subsamples are dicot in origin (17/19 and 15/17), their codon profile resembles that of the monocots in that G+C is utilized in the degenerate base III.

The use of pooled data for highly expressed genes may obscure identification of species-specific patterns in codon choice. Therefore, the codon choices of individual genes for RuBPC SSU and CAB were tabulated. The preferred codons of the maize and wheat genes for RuBPC SSU and CAB are more restricted in general than are those of the dicot species. This is in agreement with Matsuoka et al. (1987) J. Biochem. 102:673–676) who noted the extreme codon bias of the maize RuBPC SSU gene as well as two other highly expressed genes in maize leaves, CAB and phosphoenolpyruvate carboxylase. These genes almost completely avoid the use of A+T in codon position III, although this codon bias was not as pronounced in non-leaf proteins such as alcohol dehydrogenase, zein 22 kDa sub-unit, sucrose synthetase and ATP/ADP translocator. Since the wheat SSU and CAB genes have a similar pattern of codon preference, this may reflect a common monocot pattern for these highly expressed genes in leaves. The CAB gene for Lemna and the RuBPC SSU genes for Chlamdomonas share a similar extreme preference for G+C in codon position III. In dicot CAB genes, however, A+T degenerate bases are preferred by some synonymous codons (e.g., GCT for Ala, CTT for Leu, GGA and GGT for Gly). In general, the G+C preference is less pronounced for both RuBPC SSU and CAB genes in dicots than in monocots.

In designing a synthetic gene for expression in plants, attempts are also made to eliminate sequences which interfere with the efficacy of gene expression. Sequences such as the plant polyadenylation signals, e.g., AATAAA, polymerase II termination sequence, e.g., $CAN_{(7-9)}AGTNNAA$, UCUUCGG hairpins and plant consensus splice sites are highlighted and, if present in the native *Btt* coding sequence, are modified so as to eliminate potentially deleterious sequences.

Modifications in nucleotide sequence of the *Btt* coding region are also preferably made to reduce the A+T content in DNA base composition. The *Btt* coding region has an A+T content of 64%, which is about 10% higher than that found in a typical plant coding region. Since A+T-rich regions typify plant intergenic regions and plant regulatory regions, it is deemed prudent to reduce the A+T content. The synthetic *Btt* gene is designed to have an A+T content of 55%, in keeping with values usually found in plants.

Also, a single modification (to introduce guanine in lieu of adenine) at the fourth nucleotide position in the *Btt* coding sequence is made in the preferred embodiment to form a sequence consonant with that believed to function as a plant initiation sequence (Taylor et al. (1987) Mol. Gen. Genet. 210:572–577) in optimization of expression. In addition, in exemplifying this invention thirty-nine nucleotides (thirteen codons) are added to the coding region of the synthetic gene in an attempt to stabilize primary transcripts. However, it appears that equally stable transcripts are obtained in the absence of this extension polypeptide containing thirty-nine nucleotides.

Not all of the above-mentioned modifications of the natural *Bt* gene must be made in constructing a synthetic *Bt* gene in order to obtain enhanced expression. For example, a synthetic gene may be synthesized for other purposes in addition to that of achieving enhanced levels of expression. Under these conditions, the original sequence of the natural *Bt* gene may be preserved within a region of DNA corresponding to one or more, but not all, segments used to

construct the synthetic gene. Depending on the desired purpose of the gene, modification may encompass substitution of one or more, but not all, of the oligonucleotide segments used to construct the synthetic gene by a corresponding region of natural *Bt* sequence.

As is known to those skilled in the art of synthesizing genes (Mandecki et al. (1985) Proc. Natl. Acad. Sci. 82:3543–3547; Feretti et al. (1986) Proc. Natl. Acad. Sci. 83:599–603), the DNA sequence to be synthesized is divided into segment lengths which can be synthesized conveniently and without undue complication. As exemplified herein, in preparing to synthesize the *Btt* gene, the coding region is divided into thirteen segments (A–M). Each segment has unique restriction sequences at the cohesive ends. Segment A, for example, is 228 base pairs in length and is constructed from six oligonucleotide sections, each containing approximately 75 bases. Single-stranded oligonucleotides are annealed and ligated to form DNA segments. The length of the protruding cohesive ends in complementary oligonucleotide segments is four to five residues. In the strategy evolved for gene synthesis, the sites designed for the joining of oligonucleotide pieces and DNA segments are different from the restriction sites created in the gene.

In the specific embodiment, each DNA segment is cloned into a pIC-20 vector for amplification of the DNA. The nucleotide sequence of each fragment is determined at this stage by the dideoxy method using the recombinant phage DNA as templates and selected synthetic oligonucleotides as primers.

As exemplified herein and illustrated schematically in FIGS. **3** and **4**, each segment individually (e.g., segment M) is excised at the flanking restriction sites from its cloning vector and spliced into the vector containing segment A. Most often, segments are added as a paired segment instead of as a single segment to increase efficiency. Thus, the entire gene is constructed in the original plasmid harboring segment A. The nucleotide sequence of the entire gene is determined and found to correspond exactly to that shown in FIG. **1**.

In preferred embodiments the synthetic *Btt* gene is expressed in plants at an enhanced level when compared to that observed with natural *Btt* structural genes. To that end, the synthetic structural gene is combined with a promoter functional in plants, the structural gene and the promoter region being in such position and orientation with respect to each other that the structural gene can be expressed in a cell in which the promoter region is active, thereby forming a functional gene. The promoter regions include, but are not limited to, bacterial and plant promoter regions. To express the promoter region/structural gene combination, the DNA segment carrying the combination is contained by a cell. Combinations which include plant promoter regions are contained by plant cells, which, in turn, may be contained by plants or seeds. Combinations which include bacterial promoter regions are contained by bacteria, e.g., *Bt* or *E. coli*. Those in the art will recognize that expression in types of micro-organisms other than bacteria may in some circumstances be desirable and, given the present disclosure, feasible without undue experimentation.

The recombinant DNA molecule carrying a synthetic structural gene under promoter control can be introduced into plant tissue by any means known to those skilled in the art. The technique used for a given plant species or specific type of plant tissue depends on the known successful techniques. As novel means are developed for the stable insertion of foreign genes into plant cells and for manipulating the modified cells, skilled artisans will be able to select from known means to achieve a desired result. Means for introducing recombinant DNA into plant tissue include, but are not limited to, direct DNA uptake (Paszkowski, J. et al. (1984) EMBO J. 3:2717), electroporation (Fromm, M. et al. (1985) Proc. Natl. Acad. Sci. USA 82:5824), microinjection (Crossway, A. et al. (1986) Mol. Gen. Genet. 202:179), or T-DNA mediated transfer from *Agrobacterium tumefaciens* to the plant tissue. There appears to be no fundamental limitation of T-DNA transformation to the natural host range of Agrobacterium. Successful T-DNA-mediated transformation of monocots (Hooykaas-Van Slogteren, G. et al. (1984) Nature 311:763), gymnosperm (Dandekar, A. et al. (1987) Biotechnology 5:587) and algae (Ausich, R., EPO application 108,580) has been reported. Representative T-DNA vector systems are described in the following references: An, G. et al. (1985) EMBO J. 4:277; Herrera-Estrella, L. et al. (1983) Nature 303:209; Herrera-Estrella, L. et al. (1983) EMBO J. 2:987; Herrera-Estrella, L. et al. (1985) in *Plant Genetic Engineering*, New York: Cambridge University Press, p. 63. Once introduced into the plant tissue, the expression of the structural gene may be assayed by any means known to the art, and expression may be measured as mRNA transcribed or as protein synthesized. Techniques are known for the in vitro culture of plant tissue, and in a number of cases, for regeneration into whole plants. Procedures for transferring the introduced expression complex to commercially useful cultivars are known to those skilled in the art.

In one of its preferred embodiments the invention disclosed herein comprises expression in plant cells of a synthetic insecticidal structural gene under control of a plant expressible promoter, that is to say, by inserting the insecticide structural gene into T-DNA under control of a plant expressible promoter and introducing the T-DNA containing the insert into a plant cell using known means. once plant cells expressing a synthetic insecticidal structural gene under control of a plant expressible promoter are obtained, plant tissues and whole plants can be regenerated therefrom using methods and techniques well-known in the art. The regenerated plants are then reproduced by conventional means and the introduced genes can be transferred to other strains and cultivars by conventional plant breeding techniques.

The introduction and expression of the synthetic structural gene for an insecticidal protein can be used to protect a crop from infestation with common insect pests. Other uses of the invention, exploiting the properties of other insecticide structural genes introduced into other plant species will be readily apparent to those skilled in the art. The invention in principle applies to. introduction of any synthetic insecticide structural gene into any plant species into which foreign DNA (in the preferred embodiment T-DNA) can be introduced and in which said DNA can remain stably replicated. In general, these taxa presently include, but are not limited to, gymnosperms and dicotyledonous plants, such as sunflower (family Compositeae), tobacco (family Solanaceae), alfalfa, soybeans and other legumes (family Leguminoseae), cotton (family Malvaceae), and most vegetables, as well as monocotyledonous plants. A plant containing in its tissues increased levels of insecticidal protein will control less susceptible types of insect, thus providing advantage over present insecticidal uses of *Bt*. By incorporation of the insecticidal protein into the tissues of a plant, the present invention additionally provides advantage over present uses of insecticides by eliminating instances of nonuniform application and the costs of buying and applying insecticidal preparations to a field. Also, the present invention eliminates the need for careful timing of application of such preparations since small larvae are most sensitive to insecticidal protein and the protein is always present, minimizing crop damage that would otherwise result from preapplication larval foraging.

This invention combines the specific teachings of the present disclosure with a variety of techniques and expedients known in the art. The choice of expedients depends on

variables such as the choice of insecticidal protein from a *Bt* strain, the extent of modification in preferred codon usage, manipulation of sequences considered to be destabilizing to RNA or sequences prematurely terminating transcription, insertions of restriction sites within the design of the synthetic gene to allow future nucleotide modifications, addition of introns or enhancer sequences to the 5' and/or 3' ends of the synthetic structural gene, the promoter region, the host in which a promoter region/structural gene combination is expressed, and the like. As novel insecticidal proteins and toxic polypeptides are discovered, and as sequences responsible for enhanced cross-expression (expression of a foreign structural gene in a given host) are elucidated, those of ordinary skill will be able to select among those elements to produce "improved" synthetic genes for desired proteins having agronomic value. The fundamental aspect of the present invention is the ability to synthesize a novel gene coding for an insecticidal protein, designed so that the protein will be expressed at an enhanced level in plants, yet so that it will retain its inherent property of insect toxicity and retain or increase its specific insecticidal activity.

EXAMPLES

The following Examples are presented as illustrations of embodiments of the present invention. They do not limit the scope of this invention, which is determined by the claims.

The following strains were deposited with the Patent Culture Collection, Northern Regional Research Center, 1815 N. University Street, Peoria, Ill. 61604.

Strain	Deposited on	Accession #
E. coli MC1061 (p544-HindIII)	6 October 1987	NRRL B-18257
E. coli MC1061 (p544Pst-Met5)	6 October 1987	NRRL B-18258

The deposited strains are provided for the convenience of those in the art, and are not necessary to practice the present invention, which may be practiced with the present disclosure in combination with publicly available protocols, information, and materials. *E. coli* MC1061, a good host for plasmid transformations, was disclosed by Casadaban, M. J. and Cohen, S. N. (1980) J. Mol. Biol. 138:179–207.

Example 1

Design of the Synthetic Insecticidal Crystal Protein Gene

(i) Preparation of Toxic Subclones of the *Btt* Gene

Construction, isolation, and characterization of pNSB544 is disclosed by Sekar, V. et al. (1987) Proc. Natl. Acad. Sci. USA 84:7036–7040, and Sekar, V. and Adang, M. J., U.S. patent application Ser. No. 108,285, filed Oct. 13, 1987, which is hereby incorporated by reference. A 3.0 kbp HindIII fragment carrying the crystal protein gene of pNSBP544 is inserted into the HindIII site of pIC-20H (Marsh, J. L. et al. (1984) Gene 32:481–485), thereby yielding a plasmid designated p544-HindIII, which is on deposit. Expression in *E. coli* yields a 73 kDa crystal protein in addition to the 65 kDa species characteristic of the crystal protein obtained from *Btt* isolates.

A 5.9 kbp BamHI fragment carrying the crystal protein gene is removed from pNSBP544 and inserted into BamHI-linearized pIC-20H DNA. The resulting plasmid, p405/44-7, is digested with BglII and religated, thereby removing Bacillus sequences flanking the 3'-end of the crystal protein gene. The resulting plasmid, p405/54-12, is digested with PstI and religated, thereby removing Bacillus sequences flanking the 5'-end of the crystal protein and about 150 bp from the 5'-end of the crystal protein structural gene. The resulting plasmid, p405/81-4, is digested with SphI and PstI and is mixed with and ligated to a synthetic linker having the following structure:

```
          SD        MetThrAla
    5'CAGGATCCAACAATGACTGCA3'
3'GTACGTCCTAGGTTGTTACTG5' (SEQ ID NO. 5)
  SphI                  PstI
```

(SD indicates the location of a Shine-Dalgarno prokaryotic ribosome binding site.) The resulting plasmid, p544Pst-Met5, contains a structural gene encoding a protein identical to one encoded by pNSBP544 except for a deletion of the amino-terminal 47 amino acid residues. The nucleotide sequence of the *Btt* coding region in p544Pst-Met5 is presented in FIG. 1. In bioassays (Sekar and Adang, U.S. patent application Ser. No. 108,285, supra), the proteins encoded by the full-length *Btt* gene in pNSBP544 and the N-terminal deletion derivative, p544Pst-Met5, were shown to be equally toxic. All of the plasmids mentioned above have their crystal protein genes in the same orientation as the lacZ gene of the vector.

(ii) Modification of Preferred Codon Usage

Table 1 presents the frequency of codon usage for (A) dicot proteins, (B) *Bt* proteins, (C) the synthetic *Btt* gene, and (D) monocot proteins. Although some codons for a particular amino acid are utilized to approximately the same extent by both dicot and *Bt* proteins (e.g., the codons for serine), for the most part, the distribution of codon frequency varies significantly between dicot and *Bt* proteins, as illustrated in columns A and B in Table 1.

TABLE 1

		Frequency Codon Usage			
		Distribution Fraction			
Amino Acid	Codon	(A)Dicot Genes	(B)Bt Genes	(C)Synthetic Btt Gene	(D)Monocot Genes
Gly	GGG	0.12	0.08	0.13	0.21
Gly	GGA	0.38	0.53	0.37	0.17
Gly	GGT	0.33	0.24	0.34	0.18
Gly	GGC	0.16	0.16	0.16	0.43
Glu	GAG	0.51	0.13	0.52	0.75
Glu	GAA	0.49	0.87	0.48	0.25
Asp	GAT	0.58	0.68	0.56	0.27
Asp	GAC	0.42	0.32	0.44	0.73
Val	GTG	0.29	0.15	0.30	0.36
Val	GTA	0.12	0.32	0.10	0.08
Val	GTT	0.39	0.29	0.35	0.19

TABLE 1-continued

		Frequency Codon Usage			
Val	GTC	0.20	0.24	0.25	0.37
Ala	GCG	0.06	0.12	0.06	0.22
Ala	GCA	0.25	0.50	0.24	0.16
Ala	GCT	0.42	0.32	0.41	0.24
Ala	GCC	0.27	0.06	0.29	0.38
Lys	AAG	0.61	0.13	0.58	0.86
Lys	AAA	0.39	0.87	0.42	0.14
Asn	AAT	0.45	0.79	0.44	0.25
Asn	AAC	0.55	0.21	0.56	0.75
Met	ATG	1.00	1.00	1.00	1.00
Ile	ATA	0.18	0.30	0.20	0.11
Ile	ATT	0.45	0.57	0.43	0.24
Ile	ATC	0.37	0.13	0.37	0.64
Thr	ACG	0.08	0.14	0.07	0.20
Thr	ACA	0.27	0.68	0.27	0.14
Thr	ACT	0.35	0.14	0.34	0.19
Thr	ACC	0.30	0.05	0.32	0.46
Trp	TGG	1.00	1.00	1.00	1.00
End	TGA	0.33	0.00	0.00	0.34
Cys	TGT	0.44	0.33	0.33	0.30
Cys	TGC	0.56	0.67	0.67	0.70
End	TAG	0.19	0.00	0.00	0.36
End	TAA	0.48	1.00	1.00	0.30
Tyr	TAT	0.43	0.81	0.43	0.21
Tyr	TAC	0.57	0.19	0.57	0.79
Phe	TTT	0.45	0.75	0.44	0.25
Phe	TTC	0.55	0.25	0.56	0.75
Ser	AGT	0.14	0.25	0.13	0.08
Ser	AGC	0.18	0.13	0.19	0.26
Ser	TCG	0.06	0.08	0.06	0.14
Ser	TCA	0.19	0.19	0.17	0.11
Ser	TCT	0.25	0.25	0.27	0.15
Ser	TCC	0.18	0.10	0.17	0.25
Arg	AGG	0.25	0.09	0.23	0.26
Arg	AGA	0.30	0.50	0.32	0.09
Arg	CGG	0.04	0.14	0.05	0.13
Arg	CGA	0.08	0.14	0.09	0.04
Arg	CGT	0.21	0.09	0.23	0.12
Arg	CGC	0.11	0.05	0.09	0.36
Gln	CAG	0.41	0.18	0.39	0.46
Gln	CAA	0.59	0.82	0.61	0.54
His	CAT	0.54	0.90	0.50	0.33
His	CAC	0.46	0.10	0.50	0.67
Leu	TTG	0.26	0.08	0.27	0.14
Leu	TTA	0.10	0.46	0.12	0.03
Leu	CTG	0.09	0.04	0.10	0.28
Leu	CTA	0.08	0.21	0.10	0.10
Leu	CTT	0.28	0.15	0.18	0.15
Leu	CTC	0.19	0.06	0.22	0.31
Pro	CCG	0.09	0.20	0.08	0.23
Pro	CCA	0.42	0.56	0.44	0.34
Pro	CCT	0.32	0.24	0.32	0.17
Pro	CCC	0.17	0.00	0.16	0.26

154 coding sequences of dicot nuclear genes were used to compile the condon usage table. The pooled dicot coding sequences, obtained from Genbank (release 55) or, when no Genbank file name is specified, directly from the published source, were:

GENUS/SPECIES	GENBANK	PROTEIN	REF
Antirrhinum majus	AMACHS	Chalcone synthetase	
Arabidopsis thaliana	ATHADH	Alcphol dehydrogenase	
	ATHH3GA	Histone 3 gene 1	
	ATHH3GB	Histone 3 gene 2	
	ATHH4GA	Histone 4 gene 1	
	ATHLHCP1	CAB	
	ATHTUBA	α tubulin	
		5-enolpyruvyl4hyfate 3-phosphate synthetase	1
Bertholletia excelsa		High methionine storage protein	2
Brassica campestris		Acyl carrier protein	3
Brassica napus	BNANAP	Napin	
Brassica oleacea	BOLSLSGR	S-locus specific glycoprotein	
Canavalia ensiformis	CENCONA	Concanavalin A	
Carica papaya	CPAPAP	Papain	
Chlamdomonas reinhardtii	CREC552	Preapocytochrome	

TABLE 1-continued

		Frequency Codon Usage	
	CRERBCS1	RuBPC small subunit gene 1	
	CRERBCS2	RuBPC small subunit gene 2	
Cucurbita pepo	CUCPHT	Phytochrome	
Cucumis sairvus	CUSGMS	Glyoxosomal malate synthetase	
	CUSLHCPA	CAB	
	CUSSSU	RuBCP small subunit	
Daucus carota	DAREXT	Extensin	
	DAREXTR	33 kD extensin related protein	
Dolichos biflorus	DBILECS	seed lectin	
Flaveria trinervia	FTRBCR	RuBPC small subunit	
Glycine max	SOY7SAA	7S storage protein	
	SOYACT1G	Actin 1	
	SOYCHPI	CII protease inhibitor	
	SOYGLYA1A	Glycinin A1a Bx subunits	
	SOYGLYAAB	Glycinin A5A4B3 subunits	
	SOYGLYAB	Glycinin A3/b4 subunits	
	SOYGLYR	Glycinin A2B1a subunits	
	SOYHSP175	Low M W heat shock proteins	
	SOYLGBI	leghemoglobin	
	SOYLEA	Lectin	
	SOYLOX	Lipoxygenase 1	
	SOYNOD20G	20 kDa nodulin	
	SOYNOD23G	23 kDa nodulin	
	SOYNOD24H	24 kDa nodulin	
	SOYNOD26B	26 kDa nodulin	
	SOYNOD26R	26 kDa nodulin	
	SOYNOD27R	27 kDa nodulin	
	SOYNOD35M	35 kDa nodulin	
	SOYNOD75	75 kDa nodulin	
	SOYNODR1	Nodulin C51	
	SOYNODR2	Nodulin E27	
	SOYPRP1	Proline rich protein	
	SOYRUBP	RuBPC small subunit	
	SOYURA	Urease	
	SOYHSP26A	Heat shock protein 26A	
		Nuclear-encoded chloroplast heat-shock protein	4
		22 kDa nodulin	5
		β1 tubulin	6
		β2 tubulin	6
Gossypium hirsutum		Seed α globulin (vicilin)	7
		Seed β globulin (vicilin)	7
Helianthus annus	HNNRUBCS	RuBPC small subunit	
		2S albumin seed storage protein	8
Ipomoea batatas		Wound-induced catalase	9
Lemna gibba	LGIAB19	CAB	
	LGIR5BPC	RuBPC small subunit	
Lupinus luteus	LUPLBR	leghemoglobin I	
Lycopersicon esculentum	TOMBIOBR	Biotin binding protein	
	TOMETHYBR	Ethylene biosynthesis protein	
	TOMPBC2AR	Polygalacturonase-2a	
	TOMPSI	Tomato photosystem I protein	
	TOMRBCSA	RuBPC small subunit	
	TOMRBCSB	RuBPC small subunit	
	TOMRBCSC	RuBPC small subunit	
	TOMRBCSD	RuBPC small subunit	
	TOMRRD	Ripening related protein	
	TOMWIPIG	Wound induced proteinase inhibitor I	
	TOMWIPII	Wound induced proteinase inhibitor II	
		CAB 1A	10
		CAB 1B	10
		CAB 3C	10
		CAB 4	11
		CAB 5	11
Medicago sariva	ALFLB3R	Leghemoglobin III	
Mesembryanthemum crystallinum		RuBPC small subunit	12
Nicotiana plumbaginifolia	TOBATP21	Mitochondrial ATP synthase β subunit	
		Nitrate reductase	13
		Glutamine synthetase	14
nicotiana tabacum	TOBECH	Endochitinase	
	TOBGAPA	A subunit of chloroplast B3PD	

TABLE 1-continued

		Frequency Codon Usage	
	TOBGAPB	B subunit of chloroplast G3PD	
	TOBGAPC	C subunit of chloroplast G3PD	
	TOBPR1AR	Pathogenesis related protein 1a	
	TOBPR1CR	Pathogenesis-related protein 1c	
	TOBPRPR	Pathogenesis related protein 1b	
	TOBPXDLF	Peroxidase	
	TOBRBPCO	RuBPC small subunit	
	TOBTHAUR	TMV-induced protein homologous to thaumatin	
Perseus americana	AVOCEL	Cellulase	
Petroselinum hortense	PHOCHL	Chalcone synthase	
Petunia sp.	PETCA B13	CAB 13	
	PETCA B22L	CAB 22L	
	PETCA B22R	CAB 22R	
	PETCA B25	CAB 25	
	PETCA B37	CAB 37	
	PETCA B91R	CAB 91R	
	PETCHSR	Chalcone synthase	
	PETGCR1	Glycine-rich protein	
	PETRBCS08	RuBPC small subunit	
	PETRBCS11	RuBPC small subunit	
		70 kDa heat shock protein	15
Phascolus vulgaris	PHVCHM	Chintinase	
	PHVDLECA	Phytohemagglutinin E	
	PHVDLECB	Phytohemagglutinin L	
	PHVGSR1	Glutamine synthetase 1	
	PHVGSR2	Glutamine synthetase 2	
	PHVLBA	Leghemoglobin	
	PHVLECT	Lectin	
	PHVPAL	Phenylalanine ammonia lyase	
	PHVPHASAR	α phaseolin	
	PHVPHASBR	β phaseolin	
		Arcelin seed protein	16
		Chalcone synthase	17
Pisum sativum	PEAALB2	Seed albumin	
	PEACAB80	CAB	
	PEAGSR1	Glutamine synthetase (nodule)	
	PEALECA	Lectin	
	PEALEGA	Legumin	
	PEARUBPS	RuBPC small subunit	
	PEAVIC2	Vicilin	
	PEAVIC4	Vicilin	
	PEAVIC7	Vicilin	
		Alcohol dehydrogenase 1	18
		Glutamine synthetase (leaf)	19
		Glutamine synthetase (root)	19
		Histone 1	20
		Nuclear encoded chloroplast heat shock proetin	4
Raphanus satirvus		RuBPC small subunit	21
Ricinus communis	RCCAGG	Agglutinin	
	RCCRICIN	Ricin	
	RCCICL4	Isocitrate lyase	
Silene pratensis	SIPFDX	Ferrodoxin precursor	
	SIPPCY	Plastocyanin precursor	
Sinapis alba	SALGAPDH	Nuclear gene for G3PD	
Solanum tuberosum	POTPAT	Patatin	
	POTINHWI	Wound-induced proteinase inhibitor	
	POTLS1G	Light-inducible tissue specific ST-LS1 gene	
	POTP12G	Wound-induced proteinase inhibitor II	
	POTRBCS	RuBPC small subunit	
		Sucrose synthetase	22
Spinacia oleracea	SPIACPI	Acyl carrier protein I	
	SPIOEC16	16 kDa photosynthetic oxygen-evolving protein	
	SPIOEC23	23 kDa photosynthetic oxygen-evolving protein	
	SPIPCG	Plastocyanin	
	SPIPS33	33 kDa photosynthetic water oxidation complex precursor	
		Glycolate oxidase	23

TABLE 1-continued

Frequency Codon Usage			
Vicia faba	VFALBA	Leghemoglobin	
	VFALEB4	Legumin B	
		Vicillin	24

Pooled 53 monocot coding sequences obtained from Genbank (release 55) or, when no Genbank file name is specified, directly from the published source, were:

GENUS/SPECIES	GENBANK	PROTEIN	REF
Avena sativa	ASTAP3R	Phytochrome 3	
Hordeunt vulgare	BLYALR	Aleurain	
	BLYAMY1	α amylase 1	
	BLYAMY2	α amylase 2	
	BLYCHORD1	Hordein C	
	BLYGLUCB	β glucanase	
	BLYHORB	B1 hordein	
	BLYPAPI	Amylase/protease inhibitor	
	BLYTH1AR	Toxin α hordothionin	
	BLYUBIQR	Ubiquitin	
		Histone 3	25
		Leaf specific thionin 1	26
		Leaf specific thionin 2	26
		Plastocyanin	27
Oryza sativa	RICGLUTG	Glutelin	
		Glutelin	28
Triticum aestivum	WHTAMYA	α amylase	
	WHTCAB	CAB	
	WHTEMR	Em protein	
	WHTGIR	gibberellin responsive protein	
	WHTGLGB	γ gliadin	
	WHTGLIABA	α/β gliadin Class AII	
	WHTGLUT1	High MW glutenin	
	WHTH3	Histone 3	
	WHTH4091	Histone 4	
	WHTRBCB	RuBPC small subunit	
Secale cereale	RYESECGSR	γ secalin	
Zea mays	MZEA1G	40.1 kD A1 protein (NADPH-dependent reductase)	
	MZEACT1G	Actin	
	MZEADH11F	Alcohol dehydrogenase 1	
	MZEADH2NR	Alcohol dehydrogenase 2	
	MZEALD	Aldolase	
	MZEANT	ATP/ADP translocator	
	MZEEG2R	Glutelin 2	
	MZEGGST3B	Glutathione S transferase	
	MZEH3C2	Histone 3	
	MZEH4C14	Histone 4	
	MZEHSP701	70 kD Heat shock protein, exon 1	
	MZEHSP702	70 kD Heat shock protein, exon 2	
	MZELHCP	CAB	
	MZEMPL3	Lipid body surface protein L3	
	MZEPEPCR	Phosphoenolyruvate carboxylase	
	MZERBCS	RuBPC small subunit	
	MZESUSYSG	Sucrose synthetase	
	MZETP12	Triosephosphate isomerase 1	
	MZEZEA20M	19 kD zein	
	MZEZEA30M	19 kD zein	
	MZEZE15A3	15 kD zein	
	MZEZE16	16 kD zein	
	MZEZE19A	19 kD zein	
	MZEZE22A	22 kD zein	
	MZEZE22B	22 kD zein	
		Catalase 2	29
		Regulatory C1 locus	30

Bt condons were obtained from analysis of coding sequences of the following genes: Bt var. kurstaki HD-73, 6.6 kb HindIII fragment (Kronstad et al. (1983) J. Bacteriol. 154:419–428); Bt var. kurstaki HD-1, 5.3 kb fragment (Adang et al. (1987) in Biotechnology in Invertebrate Pathology and Cell Culture, K. Maramorosh (ed.), Academic Press, Inc. New York, pp. 85–99); Bt var. kurstaki HD-1, 4.5 kb fragment (Schnepf and Whiteley (1985)J. Biol. Chem. 260:6273–6280); and Bt var. tenebrionis, 3.0 kb HindIII fragment (Sekar et al. (1987) Proc. Natl. Acad. Sci. 84:7036–7040.

References

1. Klee, H. J. et al. (1987) Mol. Gen. Genet. 210:437–442.
2. Altenbach, S. B. et al. (1987) Plant Mol. Biol. 8:239–250.
3. Rose, R. E. et al. (1987) Nucl. Acids Res. 15:7197.
4. Vierling, E. et al. (1988) EMBO J. 7:575–581.
5. Sandal, N. N. et al. (1987) Nucl. Acids Res. 15:1507–1519.
6. Tingey, S. V. et al. (1987) EMBO J. 6:1–9.
7. Chlan, C. A. et al. (1987) Plant Mol. Biol. 9:533–546.
8. Allen, R. D. et al. (1987) Mol. Gen. Genet. 210:211–218.
9. Sakajo, S. et al. (1987) Eur. J. Biochem. 165:437–442.
10. Pirersky, E. et al. (1987) Plant Mol. Biol. 9:109–120.
11. Ray, J. et al. (1987) Nucl. Acids Res. 15:10587.
12. DeRocjer, E. J. et al. (1987) Nucl. Acids Res. 15:6301.
13. Calza, R. et al. (1987) Mol. Gen. Genet. 209:552–562.
14. Tingey, S. V. and Coruzzi, G. M. (1987) Plant Phys. 84:366–373.
15. Winter, J. et al. (1988) Mol. Gen. Genet. 211:315–319.
16. Osborn, T. C. et al. (1988) Science 240:207–210.
17. Ryder, T. B. et al. (1987) Mol. Gen. Genet. 210:219–233.
18. Llewellyn, D. J. et al. (1987) J. Mol. Biol. 195:115–123.
19. Tingey, S. V. et al. (1987) EMBO J. 6:1–9.
20. Gantt, J. S. and Key, J. L. (1987) Eur. J. Biochem. 166:119–125.
21. Guidet, F. and Fourcroy, P. (1988) Nucl. Acids Res. 16:2336.
22. Salanoubat, M. and Belliard, G. (1987) Gene 60:47–56.
23. Volokita, M. and Somerville, C. R. (1987) J. Biol. Chem. 262:15825–15828.
24. Bassner, R. et al. (1987) Nucl. Acids Res. 15:9609.
25. Chojecki, J. (1986) Carlsberg Res. Commun. 51:211–217.
26. Bohlmann, H. and Apel, K. (1987) Mol. Gen. Genet. 207:446–454.
27. Nielsen, P. S. and Gausing, K. (1987) FEBS Lett. 225:159–162.
28. Higuchi, W. and Fukazawa, C. (1987) Gene 55:245–253.
29. Bethards, L. A. et al. (1987) Proc. Natl. Acad. Sci. USA 84:6830–6834.
30. Paz-Ares, J. et al. (1987) EMBO J. 6:3553–3558.

For example, dicots utilize the AAG codon for lysine with a frequency of 61% and the AAA codon with a frequency of 39%. In contrast, in *Bt* proteins the lysine codons AAG and AAA are used with a frequency of 13% and 87%, respectively. It is known in the art that seldom used codons are generally detrimental to that system and must be avoided or used judiciously. Thus, in designing a synthetic gene encoding the *Btt* crystal protein, individual amino acid codons found in the original *Btt* gene are altered to reflect the codons preferred by dicot genes for a particular amino acid. However, attention is given to maintaining the overall distribution of codons for each amino acid within the coding region of the gene. For example, in the case of alanine, it can be seen from Table 1 that the codon GCA is used in *Bt* proteins with a frequency of 50%, whereas the codon GCT is the preferred codon in dicot proteins. In designing the synthetic *Btt* gene, not all codons for alanine in the original *Bt* gene are replaced by GCT; instead, only some alanine codons are changed to GCT while others are replaced with different alanine codons in an attempt to preserve the overall distribution of codons for alanine used in dicot proteins. Column C in Table 1 documents that this goal is achieved; the frequency of codon usage in dicot proteins (column A) corresponds very closely to that used in the synthetic *Btt* gene (column C).

In similar manner, a synthetic gene coding for insecticidal crystal protein can be optimized for enhanced expression in monocot plants. In Table 1, column D, is presented the frequency of codon usage of highly expressed monocot proteins.

Because of the degenerate nature of the genetic code, only part of the variation contained in a gene is expressed in this protein. It is clear that variation between degenerate base frequencies is not a neutral phenomenon since systematic codon preferences have been reported for bacterial, yeast and mammalian genes. Analysis of a large group of plant gene sequences indicates that synonymous codons are used differently by monocots and dicots. These patterns are also distinct from those reported for *E. coli*, yeast and man.

In general, the plant codon usage pattern more closely resembles that of man and other higher eukaryotes than unicellular organisms, due to the overall preference for G+C content in codon position III. Monocots in this sample share the most commonly used codon for 13 of 18 amino acids as that reported for a sample of human genes (Grantham et al. (1986 supra), although dicots favor the most commonly used human codon in only 7 of 18 amino acids.

Discussions of plant codon usage have focused on the differences between codon choice in plant nuclear genes and in chloroplasts. Chloroplasts differ from higher plants in that they encode only 30 tRNA species. Since chloroplasts have restricted their tRNA genes, the use of preferred codons by chloroplast-encoded proteins appears more extreme. However, a positive correlation has been reported between the level of isoaccepting tRNA for a given amino acid and the frequency with-which this codon is used in the chloroplast genome (Pfitzinger et al. (1987) Nucl. Acids Res. 15:1377–1386).

Our analysis of the plant genes sample confirms earlier reports that the nuclear and chloroplast genomes in plants have distinct coding strategies. The codon usage of monocots in this sample is distinct from chloroplast usage, sharing the most commonly used codon for only 1 of 18 amino acids. Dicots in this sample share the most commonly used codon of chloroplasts in only 4 of 18 amino acids. In general, the chloroplast codon profile more closely resembles that of unicellular organisms, with a strong bias towards the use of A+T in the degenerate third base.

In unicellular organisms, highly expressed genes use a smaller subset of codons than do weakly expressed genes although the codons preferred are distinct in some cases. Sharp and Li (1986) Nucl. Acids Res. 14:7734–7749 report that codon usage in 165 *E. coli* genes reveals a positive correlation between high expression and increased codon bias. Bennetzen and Hall (1982) supra have described a similar trend in codon selection in yeast. Codon usage in these highly expressed genes correlates with the abundance of isoaccepting tRNAs in both yeast and *E coli*. It has been proposed that the good fit of abundant yeast and *E. coli* mRNA codon usage to isoacceptor tRNA abundance promotes high translation levels and high steady state levels of these proteins. This strongly suggests that the potential for high levels of expression of plant genes in yeast or *E. coli* is limited by their codon usage. Hoekema et al. (1987) supra report that replacement of the 25 most favored yeast codons with rare codons in the 5' end of the highly expressed gene PGK1 leads to a decrease in both mRNA and protein. These results indicate that codon bias should be emphasized when engineering high expression of foreign genes in yeast and other systems.

(iii) Sequences within the *Btt* Coding Region having Potentially Destabilizing Influences

Analysis of the *Btt* gene reveals that the A+T content represents 64% of the DNA base composition of the coding

region. This level of A+T is about 10% higher than that found in a typical plant coding region. Most often, high A+T regions are found in intergenic regions. Also, many plant regulatory sequences are observed to be AT-rich. These observations lead to the consideration that an elevated A+T content within the *Btt* coding region may be contributing to a low expression level in plants. Consequently, in designing a synthetic *Btt* gene, the A+T content is decreased to more closely approximate the A+T levels found in plant proteins. As illustrated in Table 3, the A+T content is lowered to a level in keeping with that found in coding regions of plant nuclear genes. The synthetic *Btt* gene of this invention has an A+T content of 55%.

TABLE 3

Adenine + Thymine Content in Btt Coding Region

	Base					
Coding Region	G	A	T	C	%G+C	%A+T
Natural Btt gene	341	633	514	306	36	64
Synthetic Btt gene	392	530	483	428	45	55

In addition, the natural *Btt* gene is scanned for sequences that are potentially destabilizing to *Btt* RNA. These sequences, when identified in the original *Btt* gene, are eliminated through modification of nucleotide sequences. Included in this group of potentially destabilizing sequences are:

(a) plant polyadenylation signals (as described by Joshi (1987) Nucl. Acids Res. 15:9627–9640). In eukaryotes, the primary transcripts of nuclear genes are extensively processed (steps including 5'—capping, intron splicing, polyadenylation) to form mature and translatable mRNAs. In higher plants, polyadenylation involves endonucleolytic cleavage at the polyA site followed by the addition of several A residues to the cleaved end. The selection of the polyA site is presumed to be cis-regulated. During expression of *Bt* protein and RNA in different plants, the present inventors have observed that the polyadenylated mRNA isolated from these expression systems is not full-length but instead is truncated or degraded. Hence, in the present invention it was decided to minimize possible destabilization of RNA through elimination of potential polyadenylation signals within the coding region of the synthetic *Btt* gene. Plant polyadenylation signals including AATAAA, AATGAA, AATAAT, AATATT, GATAAA, GATAAA, and AATAAG motifs do not appear in the synthetic *Btt* gene when scanned for 0 mismatches of the sequences.

(b) polymerase II termination sequence, $CAN_{7-9}AGTNNAA$. This sequence was shown (Vankan and Filipowicz (1988) EMBO J. 7:791–799) to be next to the 3' end of the coding region of the U2 snRNA genes of *Arabidopsis thaliana* and is believed to be important for transcription termination upon 3' end processing. The synthetic *Btt* gene is devoid of this termination sequence.

(c) CUUCGG hairpins, responsible for extraordinarily stable RNA secondary structures associated with various biochemical processes (Tuerk et al. (1988) Proc. Natl. Acad. Sci. 85:1364–1368). The exceptional stability of CUUCGG hairpins suggests that they have an unusual structure and may function in organizing the proper folding of complex RNA structures. CUUCGG hairpin sequences are not found with either 0 or 1 mismatches in the *Btt* coding region.

(d) plant consensus splice sites, 5'=AAG:GTAAGT (SEQ ID NO. 6) and 3'=TTTT(Pu)TTT(Pu)T(Pu)T(Pu)T(Pu)TGCAG:C (SEQ ID NO. 7), as described by Brown et al. (1986) EMBO J. 5:2749–2758. Consensus sequences for the 5' and 3' splice junctions have been derived from 20 and 30 plant intron sequences, respectively. Although it is not likely that such potential splice sequences are present in *Bt* genes, a search was initiated for sequences resembling plant consensus splice sites in the synthetic *Btt* gene. For the 5' splice site, the closest match was with three mismatches. This gave 12 sequences of which two had G:GT. Only position 948 was changed because 1323 has the KpnI site needed for reconstruction. The 3'-splice site is not found in the synthetic *Btt* gene.

Thus, by highlighting potential RNA-destabilizing sequences, the synthetic *Btt* gene is designed to eliminate known eukaryotic regulatory sequences that affect RNA synthesis and processing.

Example 2

Chemical Synthesis of a Modified *Btt* Structural Gene

(i) Synthesis Strategy

The general plant for synthesizing linear double-stranded DNA sequences coding for the crystal protein from *Btt* is schematically simplified in FIG. 2. The optimized DNA coding sequence (FIG. 1) is divided into thirteen segments (segments A–M) to be synthesized individually, isolated and purified. As shown in FIG. 2, the general strategy begins by enzymatically joining segments A and M to form segments AM to which is added segment BL to form segment ABLM. Segment CK is then added enzymatically to make segment ABCKLM which is enlarged through addition of segments DJ, EI and RFH sequentially to give finally the total segment ABCDEFGHIJKLM, representing the entire coding region of the *Btt* gene.

FIG. 3 outlines in more detail the strategy used in combining individual DNA segments in order to effect the synthesis of a gene having unique restriction sites integrated into a defined nucleotide sequence. Each of the thirteen segments (A to M) has unique restriction sites at both ends, allowing the segment to be strategically spliced into a growing DNA polymer. Also, unique sites are placed at each end of the gene to enable easy transfer from one vector to another.

The thirteen segments (A to M) used to construct the synthetic gene vary in size. Oligonucleotide pairs of approximately 75 nucleotides each are used to construct larger segments having approximately 225 nucleotide pairs. FIG. 3 documents the number of base pairs contained within each segment and specifies the unique restriction sites bordering each segment. Also, the overall strategy to incorporate specific segments at appropriate splice sites is detailed in FIG. 3.

(ii) Preparation of Oligodeoxynucleotides

Preparation of oligodeoxynucleotides for use in the synthesis of a DNA sequence comprising a gene for *Btt* is carried out according to the general procedures described by Matteucci et al. (1981) J. Am. Chem. Soc. 103:3185–3192 and Beaucage et al. (1981) Tetrahedron Lett. 22:1859–1862. All oligonucleotides are prepared by the solid-phase phosphoramidite triester coupling approach, using an Applied Biosystems Model 380A DNA synthesizer. Deprotection and cleavage of the oligomers from the solid support are carried out according to standard procedures. Crude oligo-

nucleotide mixtures are purified using an oligonucleotide purification cartridge (OTC, Applied Biosystems) as described by McBride et al. (1988) Biotechniques 6:362–367.

5'-phosphorylation of oligonucleotides is performed with T4 polynucleotide kinase. The reaction contains 2 μg oligonucleotide and 18.2 units polynucleotide kinase (Pharmacia) in linker kinase buffer (Maniatis (1982) *Cloning Manual*, Fritsch and Sambrook (eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). The reaction is incubated at 37° C. for 1 hour.

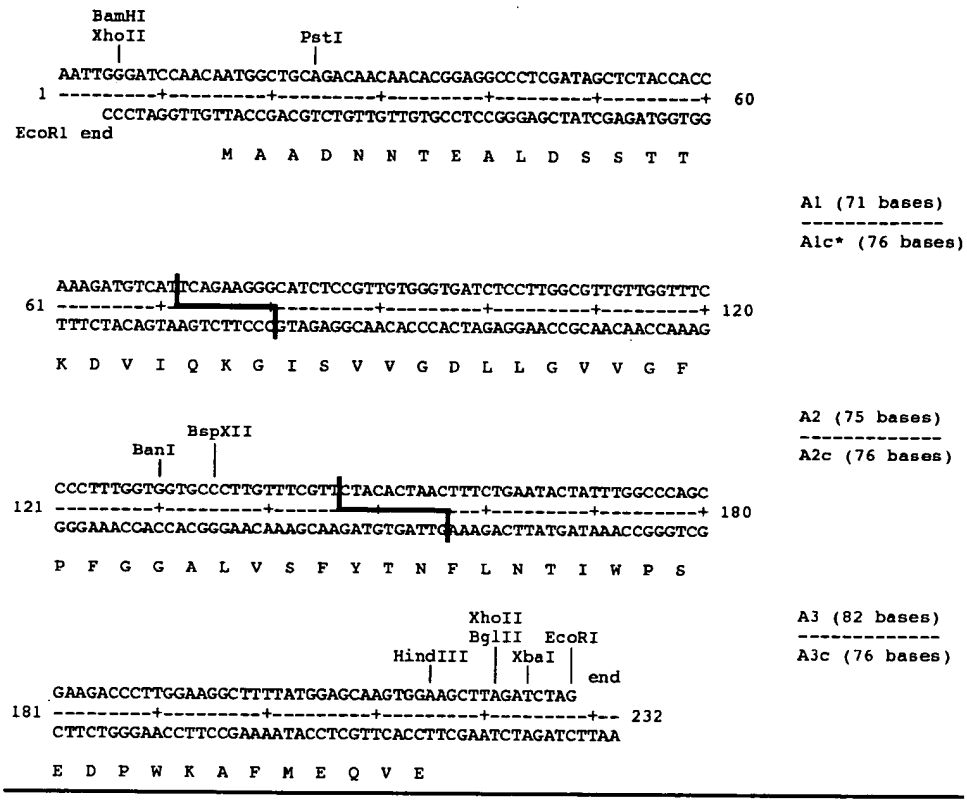
M segment DNA are used to transform *E. coli* MC1061. Colonies containing inserted blocks are identified by colony hybridization with ^{32}P-labelled oligonucleotide probes. The sequence of the DNA segment is confirmed by isolating plasmid DNA and sequencing using the dideoxy method of Sanger et al. (1977) Proc. Natl. Acad. Sci. 74:5463–5467.

(iii) Synthesis of Segment AM

Three oligonucleotide pairs (A1 and its complementary strand A1c, A2 and A2c and A3 and A3c) are assembled and ligated as described above to make up segment A. The nucleotide sequence of segment A is as follows:

TABLE 4

Nucleotide Sequence of Segment A



*c = complementary strand.

Oligonucleotides are annealed by first heating to 95° C. for 5 min. and then allowing complementary pairs to cool slowly to room temperature. Annealed pairs are reheated to 65° C., solutions are combined, cooled slowly to room temperature and kept on ice until used. The ligated mixture may be purified by electrophoresis through a 4% NuSieve agarose (FMC) gel. The band corresponding to the ligated duplex is excised, the DNA is extracted from the agarose and ethanol precipitated.

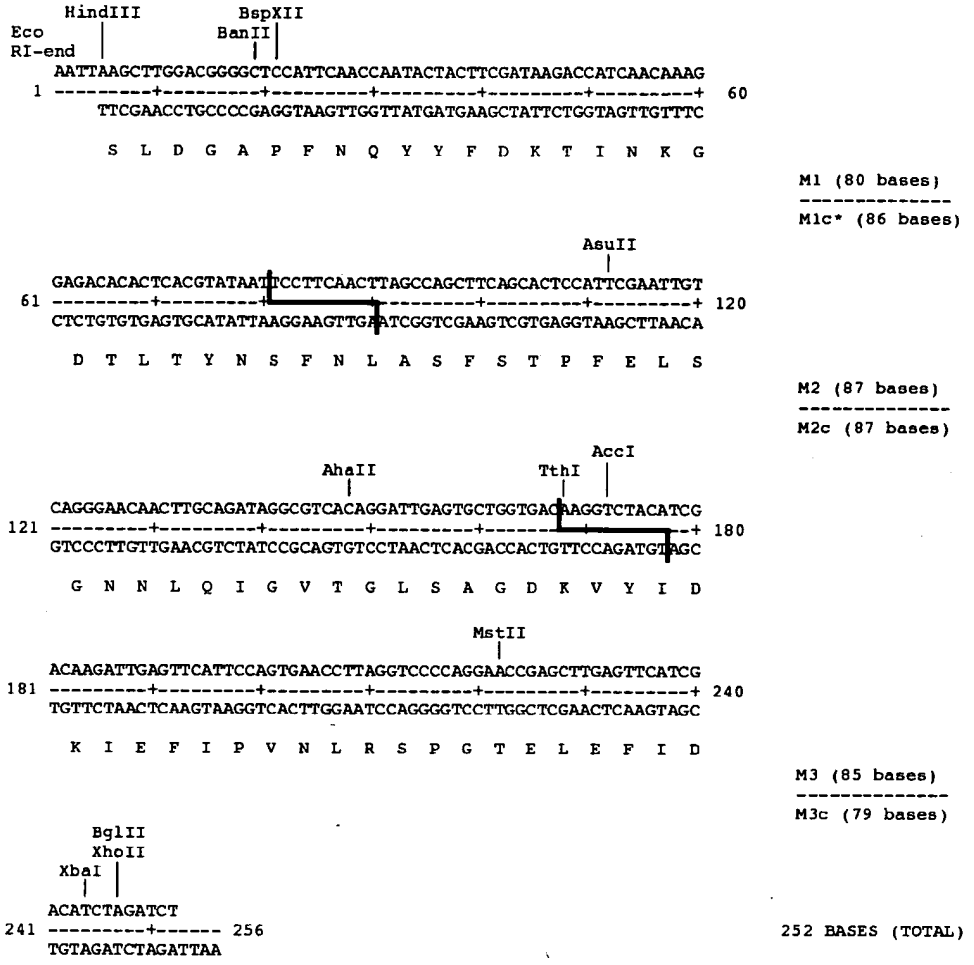
Ligations are carried out as exemplified by that used in M segment ligations. M segment DNA is brought to 65° C. for 25 min, the desired vector is added and the reaction mixture is incubated at 65° C. for 15 min. The reaction is slow cooled over 1½ hours to room temperature. ATP to 0.5 mM and 3.5 units of T4 DNA ligase salts are added and the reaction mixture is incubated for 2 hr at room temperature and then maintained overnight at 15° C. The next morning, vectors which had not been ligated to M block DNA were removed upon linearization by EcoRI digestion. Vectors ligated to the

In Table 4, bold lines demarcate the individual oligonucleotides. Fragment A1 contains 71 bases, A1c has 76 bases, A2 has 75 bases, A2c has 76 bases, A3 has 82 bases and A3c has 76 bases. In all, segment A is composed of 228 base pairs and is contained between EcoRI restriction enzyme site and one destroyed EcoRI site (5')J. (Additional restriction sites within Segment A are indicated.) The EcoRI single-stranded cohesive ends allow segment A to be annealed and then ligated to the EcoRI-cut cloning vector, pIC20K.

Segment M comprises three oligonucleotide pairs: M1, 80 bases, M1c, 86 bases, M2, 87 bases, M2c, 87 bases, M3, 85 bases and M3c 79 bases. The individual oligonucleotides are annealed and ligated according to standard procedures as described above. The overall nucleotide sequence of segment M is:

TABLE 5

Nucleotide Sequence of Segment M



*c = complementary strand

In Table 5 bold lines demarcate the individual oligonucleotides. Segment M contains 252 base pairs and has destroyed EcoRI, restriction sites at both ends. (Additional restriction sites within segment M are indicated). Segment M is inserted into vector pIC20R at an EcoRI restriction site and cloned.

As proposed in FIG. 3, segment M is joined to segment A in the plasmid in which it is contained. Segment M is excised at the flanking restrictions sites from its cloning vector and spliced into pIC20K, harboring segment A, through successive digestions with HindIII followed by BqlII. The pIC20K vector now comprises segment A joined to segment M with a HindIII site at the splice site (see FIG. 3). Plasmid pIC20K is derived from pIC20R by removing the ScaI-NdeI DNA fragment and inserting a HincII fragment containing an NPTI coding region. The resulting plasmid of *4.44* kb confers resistance to kanamycin on *E. coli*.

Example 3

Expression of Synthetic Crystal Protein Gene in Bacterial Systems

The synthetic *Btt* gene is designed so that it is expressed in the pIC20R-kan vector in which it is constructed. This expression is produced utilizing the initiation methionine of the lacZ protein of pIC20K. The wild-type *Btt* crystal protein sequence expressed in this manner has full insecticidal activity. In addition, the synthetic gene is designed to contain a BamHI site 5' proximal to the initiating methionine codon and a BglII site 3' to the terminal TAG translation stop codon. This facilitates the cloning of the insecticidal crystal protein coding region into bacterial expression vectors such as pDR540 (Russell and Bennett, 1982). Plasmid pDR540 contains the TAC promoter which allows the production of proteins including *Btt* crystal protein under controlled conditions in amounts up to 10% of the total bacterial protein. This promoter functions in many gram-negative bacteria including *E. coli* and Pseudomonas.

Production of *Bt* insecticidal crystal protein from the synthetic gene in bacteria demonstrates that the protein produced has the expected toxicity to coleopteran insects. These recombinant bacterial strains in themselves have potential value as microbial insecticides, product of the synthetic gene.

Example 4

Expression of a Synthetic Crystal Protein Gene in Plants

The synthetic *Btt* crystal protein gene is designed to facilitate cloning into the expression cassettes. These utilize sites compatible with the BamHI and BalII restriction sites flanking the synthetic gene. Cassettes are available that utilize plant promoters including CaMV 35S, CaMV 19S and the ORF 24 promoter from T-DNA. These cassettes provide the recognition signals essential for expression of proteins in plants. These cassettes are utilized in the micro Ti plasmids such as pH575. Plasmids such as pH575 containing the synthetic *Btt* gene directed by plant expression signals are utilized in disarmed *Agrobacterium tumefaciens* to introduce the synthetic gene into plant genomic DNA. This system has been described previously by Adang et al. (1987) to express *Bt* var. *kurstaki* crystal protein gene in tobacco plants. These tobacco plants were toxic to feeding tobacco hornworms.

Example 5

Assay for Insecticidal Activity

Bioassays are conducted essentially as described by Sekar, V. et al. supra. Toxicity is assessed by an estimate of the LD_{50}. Plasmids are grown in *E. coli* JM105 (Yanisch-Perron, C. et al. (1985) Gene 33:103–119). On a molar basis, no significant differences in toxicity are observed between crystal proteins encoded by p544Pst-Met5, p544-HindIII, and pNSBP544. When expressed in plants under identical conditions, cells containing protein encoded by the synthetic gene are observed to be more toxic than those containing protein encoded by the native *Btt* gene. Immunoblots ("western" blots) of cell cultures indicated that those that are more toxic have more crystal protein antigen. Improved expression of the synthetic *Btt* gene relative to that of a natural *Btt* gene is seen as the ability to quantitate specific mRNA transcripts from expression of synthetic *Btt* genes on Northern blot assays.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 9

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1794 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATGACTGCAG ATAATAATAC GGAAGCACTA GATAGCTCTA CAACAAAAGA TGTCATTCAA     60

AAAGGCATTT CCGTAGTAGG TGATCTCCTA GGCGTAGTAG GTTTCCCGTT TGGTGGAGCG    120

CTTGTTTCGT TTTATACAAA CTTTTTAAAT ACTATTTGGC CAAGTGAAGA CCCGTGGAAG    180

GCTTTTATGG AACAAGTAGA AGCATTGATG GATCAGAAAA TAGCTGATTA TGCAAAAAAT    240

AAAGCTCTTG CAGAGTTACA GGGCCTTCAA AATAATGTCG AAGATTATGT GAGTGCATTG    300

AGTTCATGGC AAAAAAATCC TGTGAGTTCA CGAAATCCAC ATAGCCAGGG GCGGATAAGA    360

GAGCTGTTTT CTCAAGCAGA AAGTCATTTT CGTAATTCAA TGCCTTCGTT TGCAATTTCT    420

GGATACGAGG TTCTATTTCT AACAACATAT GCACAAGCTG CCAACACACA TTTATTTTTA    480

CTAAAAGACG CTCAAATTTA TGGAGAAGAA TGGGGATACG AAAAAGAAGA TATTGCTGAA    540

TTTTATAAAA GACAACTAAA ACTTACGCAA GAATATACTG ACCATTGTGT CAAATGGTAT    600

AATGTTGGAT TAGATAAATT AAGAGGTTCA TCTTATGAAT CTTGGGTAAA CTTTAACCGT    660

TATCGCAGAG AGATGACATT AACAGTATTA GATTTAATTG CACTATTTCC ATTGTATGAT    720

GTTCGGCTAT ACCCAAAAGA AGTTAAAACC GAATTAACAA GAGACGTTTT AACAGATCCA    780

ATTGTCGGAG TCAACAACCT TAGGGGCTAT GGAACAACCT TCTCTAATAT AGAAAATTAT    840

ATTCGAAAAC CACATCTATT TGACTATCTG CATAGAATTC AATTTCACAC GCGGTTCCAA    900

CCAGGATATT ATGGAAATGA CTCTTTCAAT TATTGGTCCG GTAATTATGT TTCAACTAGA    960

CCAAGCATAG GATCAAATGA TATAATCACA TCTCCATTCT ATGGAAATAA ATCCAGTGAA   1020

CCTGTACAAA ATTTAGAATT TAATGGAGAA AAAGTCTATA GAGCCGTAGC AAATACAAAT   1080

CTTGCGGTCT GGCCGTCCGC TGTATATTCA GGTGTTACAA AAGTGGAATT TAGCCAATAT   1140
```

-continued

```
AATGATCAAA CAGATGAAGC AAGTACACAA ACGTACGACT CAAAAAGAAA TGTTGGCGCG   1200

GTCAGCTGGG ATTCTATCGA TCAATTGCCT CCAGAAACAA CAGATGAACC TCTAGAAAAG   1260

GGATATAGCC ATCAACTCAA TTATGTAATG TGCTTTTTAA TGCAGGGTAG TAGAGGAACA   1320

ATCCCAGTGT TAACTTGGAC ACATAAAAGT GTAGACTTTT TTAACATGAT TGATTCGAAA   1380

AAAATTACAC AACTTCCGTT AGTAAAGGCA TATAAGTTAC AATCTGGTGC TTCCGTTGTC   1440

GCAGGTCCTA GGTTTACAGG AGGAGATATC ATTCAATGCA CAGAAAATGG AAGTGCGGCA   1500

ACTATTTACG TTACACCGGA TGTGTCGTAC TCTCAAAAAT ATCGAGCTAG AATTCATTAT   1560

GCTTCTACAT CTCAGATAAC ATTTACACTC AGTTTAGACG GGGCACCATT TAATCAATAC   1620

TATTTCGATA AAACGATAAA TAAAGGAGAC ACATTAACGT ATAATTCATT TAATTTAGCA   1680

AGTTTCAGCA CACCATTCGA ATTATCAGGG AATAACTTAC AAATAGGCGT CACAGGATTA   1740

AGTGCTGGAG ATAAAGTTTA TATAGACAAA ATTGAATTTA TTCCAGTGAA TTAA         1794
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 597 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240
```

-continued

```
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
    530                 535                 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

Phe Ile Pro Val Asn
        595
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1833 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGCTGCAG ACAACAACAC GGAGGCCCTC GATAGCTCTA CCACCAAAGA TGTCATTCAG 60

AAGGGCATCT CCGTTGTGGG TGATCTCCTT GGCGTTGTTG GTTTCCCCTT TGGTGGTGCC 120

CTTGTTTCGT TCTACACTAA CTTTCTGAAT ACTATTTGGC CCAGCGAAGA CCCTTGGAAG 180

GCTTTTATGG AGCAAGTGGA AGCTTTGATG GATCAGAAGA TCGCTGATTA TGCAAAGAAC 240

AAAGCTCTTG CTGAGCTCCA GGGCCTTCAG AACAACGTCG AAGATTATGT GAGTGCACTG 300

AGTTCATGGC AAAAGAATCC TGTGTCCTCA CGAAATCCAC ATAGCCAGGG GCGCATAAGG 360

GAGCTGTTCT CTCAAGCAGA AAGTCACTTC CGGAATTCAA TGCCTTCCTT TGCCATCTCT 420

GGGTACGAGG TTCTCTTTCT TACAACCTAC GCTCAAGCTG CCAACACACA TCTGTTCTTA 480

CTAAAAGACG CTCAAATCTA TGGTGAAGAA TGGGGATACG AGAAAGAAGA TATCGCTGAG 540

TTCTACAAGC GTCAACTAAA ACTTACTCAA GAGTATACTG ACCACTGTGT CAAATGGTAT 600

AATGTTGGAT TGGATAAGTT GAGAGGTTCA TCTTATGAAT CTTGGGTAAA CTTTAACCGG 660

TACCGCAGAG AGATGACATT GACAGTGCTC GACTTGATTG CACTATTTCC ATTGTATGAT 720

GTTCGACTCT ACCCAAAGGA GGTTAAAACC GAATTGACTA GAGACGTTTT AACCGATCCC 780

ATTGTCGGAG TCAACAACCT CAGAGGCTAC GGAACAACCT TCTCTAACAT AGAAAACTAC 840

ATTCGTAAAC CACATCTATT CGACTATCTG CACAGAATTC AGTTTCACAC GCGGTTCCAA 900

CCAGGATACT ATGGAAATGA CTCTTTCAAC TATTGGTCCG GTAATTATGT TTCAACTAGA 960

CCCAGCATAG GATCTAATGA CATCATCACC TCTCCATTCT ACGGAAACAA GTCCTCCGAG 1020

CCTGTGCAAA ACTTGGAGTT TAATGGAGAG AAAGTCTATA GAGCCGTGGC CAATACCAAT 1080

CTTGCCGTCT GGCCGTCCGC TGTGTACTCA GGTGTTACCA AAGTGGAATT CAGCCAATAC 1140

AATGATCAGA CAGATGAAGC AAGTACTCAA ACTTACGACT CAAAGAGGAA TGTTGGCGCG 1200

GTCAGCTGGG ATTCTATCGA TCAACTCCCT CCAGAAACCA CCGATGAACC TCTAGAGAAG 1260

GGTTATAGCC ATCAACTCAA TTACGTAATG TGCTTTCTCA TGCAGGGTAG TAGAGGTACC 1320

ATCCCAGTGT TAACTTGGAC TCACAAGAGT GTAGACTTCT TCAACATGAT TGATTCGAAA 1380

AAGATTACTC AACTTCCGTT GGTAAAGGCC TACAAGTTAC AATCTGGTGC TTCCGTTGTC 1440

GCAGGTCCTA GGTTTACAGG AGGAGATATC ATTCAATGCA CTGAGAATGG GTCCGCGGCA 1500

ACTATCTACG TTACACCTGA TGTGTCGTAC TCTCAAAAGT ATCGTGCTAG AATTCATTAT 1560

GCTTCTACCT CTCAGATAAC ATTCACACTA AGCTTGGACG GGGCTCCATT CAACCAATAC 1620

TACTTCGATA AGACCATCAA CAAAGGAGAC ACACTCACGT ATAATTCATT CAACTTAGCC 1680

AGCTTCAGCA CTCCATTCGA ATTGTCAGGG AACAACTTGC AGATAGGCGT CACAGGATTG 1740

AGTGCTGGTG ACAAGGTTTA CATCGACAAG ATTGAGTTCA TTCCAGTGAA CCTTAGGTCC 1800

CCAGGAACCG AGCTTGAGTT CATCGACATC TAG 1833

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 610 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

-continued

```
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
```

-continued

```
        435                 440                 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
    530                 535                 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

Phe Ile Pro Val Asn Leu Arg Ser Pro Gly Thr Glu Leu Glu Phe Ile
        595                 600                 605

Asp Ile
    610
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAGGATCCAA CAATGAC 17

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAGGTAAGT 9

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TTTTUTTTUT UTUTUTGCAG C 21

-continued

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 456 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AATTGGGATC CAACAATGGC TGCAGACAAC AACACGGAGG CCCTCGATAG CTCTACCACC 60

CCCTAGGTTG TTACCGACGT CTGTTGTTGT GCCTCCGGGA GCTATCGAGA TGGTGGAAAG 120

ATGTCATTCA GAAGGGCATC TCCGTTGTGG GTGATCTCCT TGGCGTTGTT GGTTTCTTTC 180

TACAGTAAGT CTTCCCGTAG AGGCAACACC CACTAGAGGA ACCGCAACAA CCAAAGCCCT 240

TTGGTGGTGC CCTTGTTTCG TTCTACACTA ACTTTCTGAA TACTATTTGG CCCAGCGGGA 300

AACCACCACG GGAACAAAGC AAGATGTGAT TGAAAGACTT ATGATAAACC GGGTCGGAAG 360

ACCCTTGGAA GGCTTTTATG GAGCAAGTGG AAGCTTAGAT CTAGCTTCTG GGAACCTTCC 420

GAAAATACCT CGTTCACCTT CGAATCTAGA TCTTAA 456

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AATTAAGCTT GGACGGGGCT CCATTCAACC AATACTACTT CGATAAGACC ATCAACAAAG 60

TTCGAACCTG CCCCGAGGTA AGTTGGTTAT GATGAAGCTA TTCTGGTAGT TGTTTCGAGA 120

CACACTCACG TATAATTCCT TCAACTTAGC CAGCTTCAGC ACTCCATTCG AATTGTCTCT 180

GTGTGAGTGC ATATTAAGGA AGTTGAATCG GTCGAAGTCG TGAGGTAAGC TTAACACAGG 240

GAACAACTTG CAGATAGGCG TCACAGGATT GAGTGCTGGT GACAAGGTCT ACATCGGTCC 300

CTTGTTGAAC GTCTATCCGC AGTGTCCTAA CTCACGACCA CTGTTCCAGA TGTAGCACAA 360

GATTGAGTTC ATTCCAGTGA ACCTTAGGTC CCCAGGAACC GAGCTTGAGT TCATCGTGTT 420

CTAACTCAAG TAAGGTCACT TGGAATCCAG GGGTCCTTGG CTCGAACTCA AGTAGCACAT 480

CTAGATCTTG TAGATCTAGA TTAA 504

We claim:

1. A descendant plant cell comprising a pesticidal protein toxin encoded by a synthetic *Bacillus thuringiensis* (*B.t*) gene, said cell produced by the process of:

selecting a *B.t.* pesticidal protein toxin desired to be expressed in a plant cell;

obtaining a table indicating codon usage bias for a gene or genes more highly expressed in a plant cell than a native *B.t.* gene;

using said table to design a modified coding sequence which encodes said protein toxin, whereby said modified coding sequence has a frequency of codon usage that more closely resembles the frequency of codon usage of the plant cell in which it is to be expressed than did the native *B.t.* coding sequence encoding said protein toxin, said modified coding sequence having at least about 10% of the nucleotides changed as compared to the native *B.t.* coding sequence;

obtaining a synthetic *B.t.* gene comprising a coding region comprising said modified coding sequence wherein said coding region is under the control of a plant-expressible promoter;

introducing said synthetic *B.t.* gene into a plant cell;

culturing said cell to obtain descendant plant cells or plants comprising descendant plant cells, said descendant plant cells comprising said synthetic *B.t.* gene; and

establishing that said synthetic *B.t.* gene is expressed in said descendant plant cells.

2. The descendant plant cell of claim 1, wherein said modified coding sequence has an A+T base content of less than about 60%.

3. The descendant plant cell of claim 1, wherein the process of producing said plant cell contains the additional step of determining the frequency of codon usage of said modified coding sequence.

4. The descendant plant cell of claim 3, wherein said modified coding sequence comprises an A+T base content of less than about 60%.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,013,523
DATED : January 11, 2000
INVENTOR(S) : Michael J. Adang, Elizabeth E. Murray

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 6,
Line 8: "nay" should read --may--.
Line 9: Chapter 11)," should read --Chapter 1),--.

Column 14,
Line 29: "once" should read --Once--.

Column 17, Table 1,
Key: "condon" should read --codon--;
Genus/Species Column, line 1: "*Antirthinum*" should read --*Antirrhinum*--;
Protein Column, line 2: "Alcphol" should read --Alcohol--;
Genus/Species Column, line 17: "*reinhardii*" should read --*reinardtii*--.

Column 19,Table 1
Genus/Species Column, line 4: "*sairvus*" should read --*saivus*--;
Protein Column, line 6: "RuBCP" should read --RuBPC--;
Genus/Species Column, line 10: "*trincrvia*" should read --*trinvervia*--;
Genbank Column, line 13: "SOYCHPI" should read --SOYCIIPI--;
Protein Column, line 19: "leghemoglobin" should read --Leghemoglobin--;
Genus/Species Column, line 46: "*Lenuna*" should read --*Lemna*--;
Genbank Column, line 52: "TOMPBC2AR'" SHOULD READ --TOMPG2AR--;
Genus/Species Column, line 68: "*sariva*" should read --sativa--;
Genus/Species Column, line 69: "*Mesembryanthermum*" should read --*Mesembryanthemum*--;
Genus/Species Col., line 76: "*nicotiana*" should read --*Nicotiana*--;
Protein Column, line 77: "B3PD" should read --G3PD--.

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,013,523
DATED : January 11, 2000
INVENTOR(S) : Michael J. Adang, Elizabeth E. Murray

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 21,Table 1,
Genus/Species Column, line 24: "*Phascolus*" should read --*Phaseolus*--;
Genus/Species Column, line 51: "*satirvus*" should read --*sativus*--.

Column 23, Table 1,
Genus/Species Column, line 11: "*Hordeunt*" should read --*Hordeum*--.

Column 31;
Line 56: "Bq1II" should read --Bg1II--.

Column 33:
Line 5, "Ba1II" ead --Bg1II--.

Signed and Sealed this

Twenty-sixth Day of June, 2001

Nicholas P. Godici

Attest:

Attesting Officer

NICHOLAS P. GODICI
Acting Director of the United States Patent and Trademark Office